The High Redshift Universe: Galaxies and the Intergalactic Medium

Koki Kakiichi

München 2016

The High Redshift Universe: Galaxies and the Intergalactic Medium

Koki Kakiichi

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Koki Kakiichi
aus Komono, Mie, Japan

München, den 15 Juni 2016

Erstgutachter: Prof. Dr. Simon White
Zweitgutachter: Prof. Dr. Jochen Weller
Tag der mündlichen Prüfung: Juli 2016

Contents

List of Figures

List of Tables

Zusammenfassung

Diese Dissertation behandelt die Natur der Epoche der Reionisierung (Epoch of Reionization, EoR) und das Zusammenspiel von hoch rot-verschobenen Galaxien und Quasaren mit dem intergalaktischen Medium im Universum. Das Verstndniss der EoR wird als eines der wichtigsten Probleme der extragalaktischen Astronomie und Kosmologie des 21. Jahrhunderts angesehen. Trotz ihrer Wichtigkeit ist die Geschichte und Morphologie der Reionisierung und der fr sie verantwortlichen Quellen (Galaxien und/oder Quasare) weiterhin unbekannt. In der vorliegenden Arbeit versuche ich (1) die Geschichte der Reionisierung mithilfe von kosmologischen Strahlungstransportsimulationen und Durchmusterungen von Lyα emittierender Galaxien bei hohen Rotverschiebungen einzugrenzen, (2) die physikalischen Prozesse der durch Galaxien und Quasare vorangetriebenen Reionisierung zu untersuchen und (3) die messbaren Radiosignale aus der Umgebung von Quasaren zu studieren. Diese Studien werden als Trittsteine fuer die Interpretation und Analyse von ultra-tiefen Galaxiendurchmusterungen und zuknftigen interferrometrischen Radiobeobachtungen dienen, um mithilfe der 21-cm-Kosmologie Licht auf die Geschichte der EoR zu werfen.

Im ersten Teil werde ich die Durchmusterungen von hoch-rotverschobenen Lyα emmittierenden Galaxien mithilfe von kosmologisch Strahlungstransportsimulationen interpretieren um die Geschichte der Reionisierung einzugrenzen. Eine Neuerung in dieser Studie ist die Analyse von beobachteten Eigenschaften von Lyα-emmittierenden Galaxien mithilfe der Simulation eines groen Parameterraums, der sich ber verschiedene groskalige Reionisierungsmorphologien und die kleinskaligen selbstabschirmenden Strukturen des IGM, wie etwa Lyman-Limit-Systeme und gedaempfte Lyα-Systeme, erstreckt. Durch den Vergleich dieser Modelle mit den beobachteten Lyα-Leuchtkraftfunktionen bei Rotverschiebung $z \approx 7$ erhalte ich eine neue Grenze fuer den Anteil neutralen Wasserstoffs bei $z \approx 7$ in der Grenordnung von einigen zehn Prozent, allerdings mit groen Unsicherheiten in Modell und Beobachtung. Zur Erklrung dieser Daten wird der groe neutrale Anteil von $\gtrsim 1\%$ bevorzugt, welcher relativ wenig von der groskaligen Reionisierungsmorphologie und den kleinskaligen selbstabschirmenden Strukturen des IGM beeinflusst wird. Zusammen mit der Einschrnkung der Gunn-Peterson'schen optischen Dicke durch Quasarspektren deuten diese neuen Erkenntnisse auf eine spte und schnelle Reionisierungsgeschichte hin, die auch mit den Werten der Planck-Mission aus dem kosmischen Mikrowellenhintergrund vereinbar ist.

Im zweiten Teil studiere ich die Physik der Reionisierungsprozesse in Quasarumgebungen bei hoher Rotverschiebung mithilfe eines groen Satzes von Multifrequenz Strahlungstrans-

portsimulationen im UV- und Rntgenspektrum. Diese detailllierte Untersuchung hilft zu verstehenen, wie Galaxien und Quasare gemeinsam die Wasserstoff- und Heliumreionisierung vorantreiben und den Zustand des IGM beeinflussen. Ich zeige, dass in Quasarumgebungen die zentralen Quasare und die sie umgebenden Galaxien die Reionisierungsmorphologie von Wasserstoff und Helium, sowie den thermischen Zustand des IGM gemeinsam beeinflussen. Die Wasserstoff Reionisierungsmorphologie ist das Ergebnis der berlagerung des Einflusses von Galaxien und Quasaren. Es wird daher mglich sein, wenn auch mit Schwierigkeiten, Aussagen ber die Quellen der Reionisierung alleine aufgrund der Reionisierung von Wasserstoff zu treffen. Andererseits enthlt die Heliumreionisierungsmorphologie und die thermische Struktur des IGM klare Hinweise auf die Rolle von Quasaren waehrend der Reionisierung. Beobachtungen des intergalaktischen Heliums und der Temperatur des IGM, zusammen mit herkmmlichen Beobachtungen der Wasserstoffreionisierung werden Aufschluss ber die Rolle von Galaxien und Quasaren im Vorantreiben der Reionisierung geben.

Beobachtbare Signaturen der Quasarumgebungen bei hohen Rotverschiebungen werden im dritten Teil untersucht. Die H I 21cm Signale aus den Umgebungen der Quasare ermglichen die direkte Beobachtung des Wasserstoffreionisierungsprozesses, whrend das Signal der ^{3}He II 3.46cm Linie des intergalaktischen Heliums komplementre Informationen ber den Beitrag von Quasaren liefert. Beobachtungen dieser Radiosignale knnen mit in Planung befindlichen Radiointerferometern wie dem Square Kilomter Array realisiert werden. Ich schlage vor mithilfe von Galaxiendurchmusterungen in Bereichen des Himmels, fr die 21cm-Tomographien vorliegen direkt die Rolle von Galaxien und Quasaren zu untersuchen.

Zusammenfassend kann gesagt werden, dass zuknftige Galaxiendurchmusterungen im Optischen und Infraroten und 21cm-Experimente mit Radiointerferometern unser Verstndnis der EoR stark erweitern werden. Diese Beobachtungen profitieren von kosmologischen, hydrodynamischen Strahlungstransportsimulationen, die es uns erlauben die komplexen Daten zu interpretieren und zu analysieren und physikalische Einschrnkungen fr die EoR zu finden. Das Herzstck des Reionisierungsproblems ist das Zusammenspiel von Galaxien und Quasaren mit dem IGM. In dieser Hinsicht argumentiere ich auf Grundlage vorliegender Dissertation, dass mithilfe von spektroskopischen Untersuchungen von Lyα-Galaxien in Quasarfeldern mit hoher Rotverschiebung in naher Zukunft Fortschritte erzielt werden, da hier sowohl die Eigenschaften der frhen Galaxien als auch des IGM direkt untersucht werden knnen. Auerdem erlaubt es eine solche Galaxiendurchmusterungsstrategie einzigartige Bereiche des Himmels zur vollen Entfaltung des Potentials der 21cm-Kosmologie zu nutzen.

Summary

This thesis concerns the nature of the Epoch of Reionization (EoR) and the interplay between the high-redshift galaxies/quasars and the intergalactic medium (IGM) in the universe. Understanding the EoR is considered as one of the most important problems in extragalactic astronomy and cosmology in the 21st century. However, the history and the morphology of the reionization process and the ionizing sources (galaxies and/or quasars) are still unknown. In this thesis, I attempt (1) to constrain the reionization history using cosmological radiative transfer simulations and high-redshift surveys of Lyα emitting galaxies, (2) to study the physics of reionization driven by galaxies and quasars, and (3) to study the observational radio signals from quasar environments. These studies will serve as stepping stones for interpreting and analysing ultra deep galaxy surveys as well as upcoming radio interferometric observations for 21cm cosmology to shed light on the nature of EoR.

In the first part, I interpret the high redshift surveys of Lyα emitting galaxies using a suite of cosmological radiative transfer simulations to derive a constraint on the reionization history. A novelty of this study is to analyse the observed properties of Lyα emitting galaxies using a large model grid spanning different large-scale reionization morphology and the small-scale self-shielded structures of the IGM such as Lyman limit systems and damped Lyα systems. Comparing these models to the observed Lyα luminosity function at redshift $z \sim 7$, I derive a new constraint on the neutral hydrogen fraction at $z \sim 7$ to be of order of tens of per cent, albeit with large observational and theoretical uncertainties. In order to explain the data, the large neutral hydrogen fraction of $\gtrsim 1\%$ is favoured, which is relatively insensitive to the detailed large-scale reionization morphology and the small-scale self-shielded structures of the IGM. Together with the Gunn-Peterson optical depth constraint from the quasar spectra, this new constraint suggests a late and rapid reionization history, which is also consistent with the CMB constraint from Planck.

In the second part, I study the physics of the reionization process in high redshift quasar environments using a large suite of multi-frequency (UV and X-ray) radiative transfer simulations. This detailed study allows us to understand how galaxies and quasars collectively drive hydrogen and helium reionization, and impact the physical state of the IGM. I show that in quasar environments both the central quasars and the surrounding galaxies concertedly control the reionization morphology of hydrogen and helium and nonlinearly impact the thermal structure of the IGM. The hydrogen reionization morphology is a cumulative result of both galaxies and quasars. Therefore, it will be difficult, although possible, to

learn about the source of reionization by probing of hydrogen reionization alone. However, the helium reionization morphology and thermal structure of the IGM contain tell-tale signs to understand the role of quasars during reionization. Observational probes of the intergalactic helium and temperature of the IGM, together with canonical probes for hydrogen reionization, will shed light on the role of galaxies and quasars in driving reionization.

Observational signatures of the high redshift quasar environments are studied in the third part. The H I 21cm signal from the quasar environments probe directly the reionization process of hydrogen. Furthermore, the ^{3}He II 3.46cm signal from the intergalactic helium provides a complementary information about the contribution of quasars to reionization. The observations of these radio signals can be realized by planned radio interferometers such as the Square Kilometer Array. I suggest that a direct test to understand the role of galaxies and quasars during reionization is to perform galaxy surveys in a region of sky imaged by 21cm tomography.

In conclusion, future galaxy surveys using optical/infrared facilities and 21cm experiments by radio interferometers will greatly advance our understanding of the nature of the EoR. Such observations will benefit from cosmological hydrodynamical, radiative transfer simulations, which allow us to interpret and analyse the complex data to derive physical constraints on the EoR. The heart of the reionization problem lies in the interplay between galaxies/quasars and the IGM. In this regard, based on the work presented in the thesis, I argue that the near-future advance will come by spectroscopic observations of Lyα galaxies in high redshift quasar fields, by which both the properties of early galaxies and the IGM can be directly studied using the galaxies' and the background quasar spectra. In addition, such galaxy survey strategy offers unique fields on the sky for future radio interferometric observations to utilise the full potential of 21cm cosmology.

Chapter 1

Extragalactic Astrophysics and Cosmology

1.1 Prologue

Since the time of Galileo[162], our endeavor to understand the nature of physical world – the universe – begins with observing the night sky, from which we attempt to make sense of all astronomical phenomena in a coherent systematic picture. The intricate, yet rich and diverse range of phenomena from stars to galaxies and the large-scale structure of the universe that cannot be imagined from seemingly simple underlying physical principles adds colours and lives to our picture of the world.

In the 21st century, building upon the breakthrough during the last century, we have established the Standard Model of cosmology (ΛCDM cosmology). In this picture, the universe, which consists of dark matter, dark energy, and ordinary baryonic matter, has grown from the phase of hot and dense state – hot big bang – seeded with the quantum fluctuation originated from earlier inflationary epoch. Since then, the universe has expanded and structures such as galaxies have formed out of the growth of the relic quantum fluctuations amplified by the gravitational instability.

On the landscape of cosmological model, extragalactic astrophysics sets arena for rich astronomical phenomena. Here, almost infinitely many questions concerning our origins and our place in the universe can be asked. How did our home Milky Way galaxy emerge? Why is there such a diverse range of galaxies seen in the sky? What is the origin of galaxies and their distribution in the universe? All questions about galaxies and their environments are fundamental for our understanding of the physical universe. They are like beautiful flowers and forests in a massive mountain range, which add colours and lives in our view of the world. What I would like to tackle in this thesis is a question concerning the early history – the era known as *Cosmic Dawn* – when stars and galaxies are formed for the first time in the universe. This problem is regarded as one of the most important outstanding mysteries in the nature of the physical world in the 21st century astronomy. The frontier of this field is well captured by the Hubble Ultra Deep Field image (Figure 1.1).

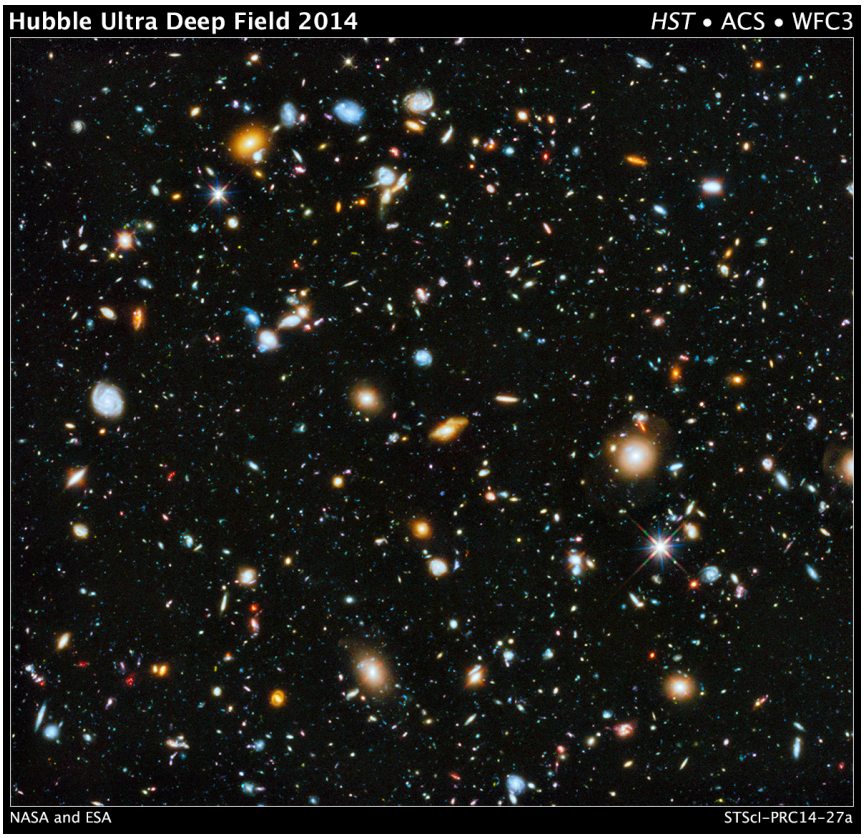

Figure 1.1: The Hubble Ultra Deep Field 2014 image. (Image Credit: NASA, ESA, H. Teplitz and M. Rafelski (IPAC/Caltech), A. Koekemoer (STScI), R. Windhorst (Arizona State University), and Z. Levay (STScI))

Problems of Cosmic Dawn are closely linked with that of the Epoch of Reionization. Galaxies consist of only several per cent of the total baryonic matter in the universe. The rest of baryons resides in the space between galaxies in the form of diffuse gas, known as intergalactic medium (IGM). Therefore, it is better pictured that galaxies reside in the vast ocean of the intergalactic gas. The epoch of reionization is the era when the intergalactic gas has drastically transformed its physical state from neutral to ionized. Current consensus is that reionization is driven by the ionizing radiation from the early descendants of first galaxies or black holes, although one would agree that the truth of the reionization epoch is still unknown. In this regard, reionization is fundamentally a problem of the interplay between early galaxies and the intergalactic medium. Understanding the Epoch of Reionization is therefore a key to understand the era of Cosmic Dawn.

This chapter is largely devoted to build a good anchor to start a voyage into the high-redshift $z > 6$ universe, without which we may get lost in the sea of unknowns. For this reason, I discuss, in a large fraction of this chapter, about galaxy formation and the IGM at $2 < z < 5$.

1.2 Briefly Story about Reionization

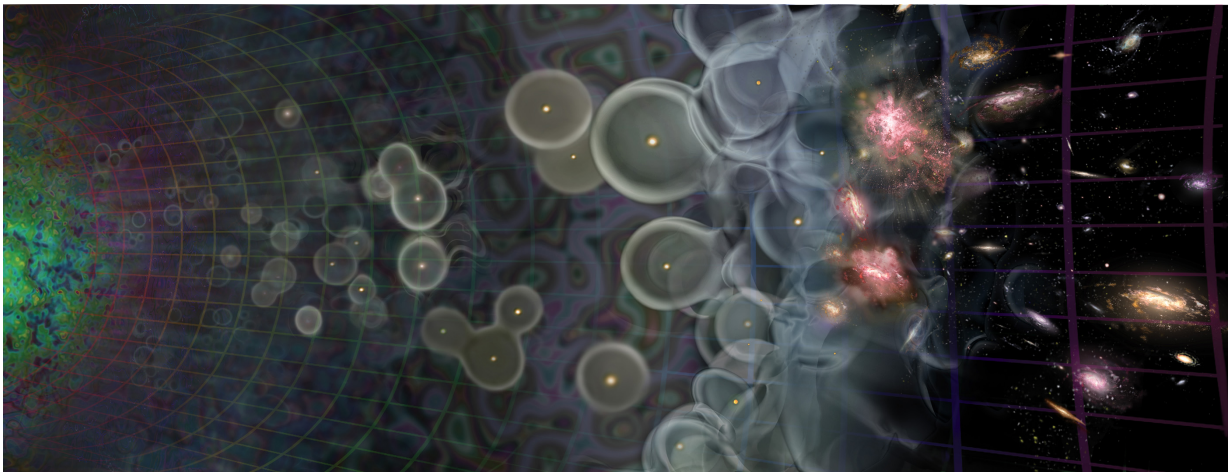

Figure 1.2: Cartoon illustration of the Epoch of Reionization and cosmic dawn. Credit: Loeb (2006) Scientific American

A key frontier of cosmology is to understand the first billion years of cosmic history (e.g. [255]). During this era, the universe has undergone two landmark events known as Cosmic Dawn and the Epoch of Reionization (EoR). Cosmic Dawn is the era when stars, galaxies, and black holes are formed for the first time in cosmic history, ending the Dark Age when there were no luminous objects in the universe. These first galaxies and their early descendants are thought to be responsible for transforming the entire cosmic environment from cold and neutral gas of primordial hydrogen and helium to ionized gas as in the present day universe. This era of cosmic phase transition is known as the Epoch of Reionization. The cartoon picture in Figure 1.2 illustrates our conception of how this process occurred. Early galaxies at redshifts $z \sim 6 - 25$ ionize the surrounding medium creating large H II regions. These cosmological H II regions gradually percolate the entire intergalactic gas in the universe, producing even larger ones by overlapping of many H II regions. Then the reionization process is completed when there is no primordial neutral gas to be reionized. This picture has been proposed even in the 90's. It is a goal of 21st century astronomy to test this paradigm and construct a true story about Cosmic Dawn and EoR based on solid observations.

1.3 Foundation of Observational Cosmology

Historically, observations of Slipher and Hubble had played a pivotal role in establishing the modern picture that the universe is expanding, which had transformed a traditionally held conception of static universe before 1920's (see [340] for historical account). In the period of 1914-1917, Slipher first found the tendency that galaxies' spectra are systematically redshifted ([410]), which can be interpreted in favour of a cosmological expansion model.

Later, together with a theoretical motive for an expanding universe model proposed by de Sitter and others, Lundmark and Hubble searched for the linear redshift-distance relation to provide a further observational evidence for the expanding universe model ([257, 201, 202]). Today, modern distance measurements using Cepheid and Supernova, together with other probes (§ 1.5), place the expanding universe model on a firm observational ground.

The cosmological principle states that the universe is both homogeneous and isotropic on the largest scale. This principle is supported by the modern observational evidence for the isotropy of the cosmic microwave background (CMB) and the homogeneity of the distribution of galaxies on very large scales (e.g. [454]). An alternative argument to arrive at cosmological principle is to combine the observed isotropy of the sky and Copernican principle[1]. We can prove that, by a geometrical argument, the former two imply the homogeneity of matter distribution in the universe ([339]).

To chart galaxies in the universe, firstly we should lay out the metric $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$ where ds is the spacetime interval, $g_{\mu\nu}$ is the metric tensor, and x^μ is the spacetime coordinate. Only on the basis of symmetry ([454], see also Peter Scheuer's argument [256]), the cosmological principle constrains the form of metric to be the Robertson-Walker (RW) metric. This can be written with a spherical coordinate (r, θ, ϕ) around an observer as,

$$ds^2 = c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right], \tag{1.1}$$

where t is the cosmological time and $a(t)$ is the scale factor. The physical radial interval between two events is $dl_p = a(t)dr/\sqrt{1-kr^2}$, where the comoving coordinate r conveniently factored out the effect of cosmological expansion. k parameterizes the three spatial geometries compatible with the cosmological principle. The universe could be spatially flat ($k = 0$), closed ($k = 1$), or open ($k = -1$).

The RW metric allows us to connect the observed redshift of galaxies to their spacetime positions in the universe. As an observational fact, we see the redshifted light from distant galaxies. This defines the redshift $1 + z \equiv \lambda_{\rm obs}/\lambda_{\rm em}$ where $\lambda_{\rm obs}$ is the observed wavelength and $\lambda_{\rm em}$ is the rest-frame wavelength emitted from a galaxy. The light from a galaxy travels towards us following the null geodesics $cdt = a(t)dr/\sqrt{1-kr^2}$. For a galaxy emitting at time $t_{\rm em}$ and at radial distance r away from us, the two wave-fronts of light separated by a time interval $\Delta t_{\rm em}$, i.e wavelength $\lambda_{\rm em} = c\Delta t_{\rm em}$, are received by an observer at time $t_{\rm obs}$ and $t_{\rm obs} + \Delta t_{\rm obs}$ as wavelength $\lambda_{\rm obs} = c\Delta t_{\rm obs}$. In the universe described by the RW metric, the emitted and observed time of the first wave-front is related to the radial distance to the galaxy,

$$\int_{t_{\rm em}}^{t_{\rm obs}} \frac{cdt}{a(t)} = \int_0^r \frac{dr}{\sqrt{1-kr^2}}. \tag{1.2}$$

Because the radial distance to the galaxy (the right-hand-side of the above formula) must be unchanged between the emission time interval of the first and second wave-fronts, we have $\int_{t_{\rm em}}^{t_{\rm obs}} \frac{cdt}{a(t)} = \int_{t_{\rm em}+\Delta t_{\rm em}}^{t_{\rm obs}+\Delta t_{\rm obs}} \frac{cdt}{a(t)}$, thus $\lambda_{\rm obs}/\lambda_{\rm em} = 1 + z = a(t_{\rm obs})/a(t_{\rm em})$. For the scale factor

[1] Copernican principle states that we are not occupying a special place in the universe

normalized at the present day, the redshift of a galaxy is related to the cosmological time coordinate of the object through $1+z = 1/a(t)$ regardless of the geometry of the universe. This one-to-one correspondence between redshift and scale factor (time) allows us to use redshift as a convenient label to mark the time coordinate of galaxies. This interpretation of redshift in terms of cosmological expansion is the basis for mapping the redshifts of galaxies to the 3D distribution of galaxies in the universe.

The above two fundamental concepts (RW metric and cosmological redshift) do not rely on a specific theory of gravity such as general relativity, but only on the cosmological principle and the geometrization of spacetime. Physical assumptions made so far are the homogeneity (or Copernican principle), the isotropy of matter distribution on large scales, and the invariance of speed of light. To describe the expansion history of the universe, we finally need to specify the law of gravity describing the dynamics of spacetime. In general relativity, the Einstein field equation describes dynamics of a metric tensor $g_{\mu\nu}$,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.3}$$

where the Einstein tensor $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$ is given by the Ricci tensor $R_{\mu\nu}$ and Ricci scalar R, and Λ is the cosmological constant. The total density ρ and the pressure p in the universe contribute to the dynamics through the energy-momentum tensor $T_{\mu\nu} = (\rho + pc^{-2}) U_\mu U_\nu + p g_{\mu\nu}$ where U^μ is the four-velocity. The Friedmann equation describing the expansion rate of the universe, i.e. Hubble parameter $H(z)$, is derived by substituting the RW metric into the Einstein field equation (see [127] for detail),

$$H(z) = H_0 [\Omega_\Lambda + \Omega_k (1+z)^2 + \Omega_m (1+z)^3 + \Omega_r (1+z)^4]^{1/2}. \tag{1.4}$$

Here the density parameter $\Omega_i = \rho_i(0)/\rho_{crit}(0)$ is defined in terms of the present day critical density $\rho_{\rm crit}(0) = 3H_0^2/(8\pi G)$. We write Ω_Λ for dark energy, Ω_m for total (dark matter and baryon), Ω_r for radiation, and $\Omega_k = 1 - \Omega_\Lambda - \Omega_m - \Omega_r$ for curvature (total density). $H_0 = 100h$ km s^{-1} Mpc^{-1} is the present day expansion rate where h is the dimensionless Hubble parameter. The total matter content, dark energy, and the geometry of the universe are closely related to the expansion history, which in turn is essential for extragalactic astronomy to chart out the distribution of galaxies and the intergalactic gas in the universe.

1.4 Hierarchical Structure Formation

The large-scale structure in the universe is thought to be formed out of the small fluctuations originated from the earlier inflationary epoch. A concordance model of inflation predicts that this tiny density perturbation is an almost Gaussian random field with a primordial power-law power spectrum $P(k) \propto k^{n_s}$ having a primordial power-law index $n_s \approx 1$. This density perturbation is amplified by the action of gravitational instability through various cosmological epochs from the radiation-dominated era at $z \gtrsim 2.4 \times 10^4 \Omega_m h^2 \simeq 3500$,

and the matter-dominated era at $0.5 \lesssim z \lesssim 3500$, to the dark energy-dominated era at $z \lesssim (\Omega_\Lambda/\Omega_m)^{1/3} - 1 \simeq 0.5$. The linear growth of the perturbations at each epoch is summarized in the transfer function $T(k)$. In the standard cosmological paradigm, the fluctuations in the cosmic microwave background and the large-scale distribution of galaxies seen in the sky (see Figure 1.3) are the striking manifestation of this gravitational growth of the relic quantum fluctuation.

I discuss the structure formation from the matter-dominated era to the present day. I focus on the redshift range of $2 \lesssim z \lesssim \mathcal{O}(10^2)$ (i.e. after the redshift of the last-scattering surface of the cosmic microwave background at $z \simeq 1100$ when the ionized plasma in the universe recombined to neutral) because the growth of structure during this epoch is the most relevant for understanding Cosmic Dawn and the Epoch of Reionization.

In this epoch, the growth of structure is dominated by the gravitational field of cold dark matter, which can be formulated as the collisionless gravitational dynamics in a cosmological background. Mathematically, it is described by the Boltzmann equation and general relativity (see [27] and Chapter 4 of [311] for more detail). On the scales of our interest, the Newtonian limit of gravity in the expanding universe is a good approximation. In terms of the comoving coordinates $\boldsymbol{r}$, the Lagrangian of a particle of mass m in an expanding universe is $L = (1/2)mv^2 - m\phi$ where $\boldsymbol{v} = a\dot{\boldsymbol{r}}$ is the peculiar velocity and ϕ is the gravitational potential fluctuation. The evolution of a system is then described by the distribution function in phase space $f(\boldsymbol{r}, \boldsymbol{p}, t)$, which follows the collisionless Boltzmann equation as a direct consequence of the conservation of the number of particles,

$$\frac{\partial f}{\partial t} + \frac{\boldsymbol{p}}{ma^2} \cdot \frac{\partial f}{\partial \boldsymbol{r}} - m\boldsymbol{\nabla}\phi \cdot \frac{\partial f}{\partial \boldsymbol{p}} = 0, \tag{1.5}$$

where $\boldsymbol{p} \equiv \partial L/\partial \dot{\boldsymbol{r}} = ma\boldsymbol{v}$ is the canonical momentum, and the gravitational potential fluctuation is given by the Poisson's equation,

$$\nabla^2\phi = 4\pi G\bar{\rho}a^2\delta. \tag{1.6}$$

This Poisson's equation is the Newtonian limit to general relativity, which is sourced by the density fluctuation δ where $\bar{\rho} = \bar{\rho}(0)a^{-3}$ is the proper physical density. Because of the obvious difficulty to solve the Boltzmann equation directly, there are three primary techniques to seek the properties of the solution: (1) perturbation theory, (2) Press-Schechter and peak formalism, and (3) N-body simulations.

In perturbation theory, we take the velocity moments of the distribution function. This reduces the Boltzmann equation to a fluid description, which is valid in the single stream regions before shell crossing. Then, we employ the perturbative approach in overdensity δ. At early times when the perturbation is small, $\delta < 1$, the evolution of the overdensity is well described by the linear theory, after applying the above procedure,

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\bar{\rho}\delta. \tag{1.7}$$

In the linear regime, the perturbation simply grows its amplitude as $\delta \propto D(z)$ according to the linear growth rate D (e.g. [243]),

$$D(z) \propto H(z) \int_z^\infty \frac{1+z'}{H^3(z')} dz'. \tag{1.8}$$

In the matter-dominated era, the linear perturbation grows as the scale factor, $D(z) \propto a(z)$ without any mixing of Fourier modes. Therefore the matter power spectrum simply changes its normalization in the linear regime. The linear matter power spectrum is

$$P(k, z) \propto k^{n_s} D^2(z) T^2(k), \tag{1.9}$$

where $T(k)$ is the transfer function encoding the impact of early time phenomena $z \lesssim 1100$ on the power spectrum. It is convenient to write in terms of the dimensionless power spectrum $\Delta^2(k, z) = k^3 P(k, z)/(2\pi^2)$, which is the contribution to the rms fluctuation per unit logarithmic bin at k.

When the fluctuation grows to $\delta > 1$, the gravitational collapse leads to the formation of virialised objects, called dark matter haloes. According to the spherical top-hat model, the collapse corresponds to the time when a (extrapolated) linear perturbation reaches a critical value $\delta_c \simeq 1.69$. Properties of a halo approximately follow the virial theorem according to which the gravitational energy is sustained by the internal velocity dispersion of dark matter. At virialization, the mean overdensity becomes $\Delta_c \approx 178 \sim 200$ in a spherical collapse model. The virial radius of a halo of mass M_h follows from the definition $M_h = (4/3)\pi \Delta_c \bar{\rho}_m R_{\rm vir}^3$,

$$R_{\rm vir} = 52.3 \left(\frac{M_h}{10^{10}\ {\rm M}_\odot} \right)^{1/3} \left(\frac{\Omega_m}{0.3} \frac{\Delta_c}{200} \right)^{-1/3} h^{-1}{\rm ckpc}. \tag{1.10}$$

The formation and clustering of haloes are of our central interest because in the standard paradigm the formation of galaxies occurs inside haloes and the large-scale distribution of galaxies is related to the halo clustering.

The extended Press-Schetchter formalism provides us a useful analytic framework to understand the properties of haloes such as the mass function, merger rate, and clustering. The idea in the Press-Schechter formalism ([353]) is drawn from linear theory and spherical top-hat collapse. It relates the fraction of overdensity greater than $\delta_c \simeq 1.69$ smoothed on a filtering scale $R_f = (3M_h/4\pi\bar{\rho}_m(0))^{1/3}$ to the number of haloes with mass M_h. The halo mass function is then given by

$$\frac{dn}{d\ln M_h} = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}_m(0)}{M_h} \nu e^{-\frac{\nu^2}{2}} \left| \frac{d\ln\sigma(M_h, z)}{d\ln M} \right| \tag{1.11}$$

where $\nu = \delta_c/\sigma(M_h, z)$ and $\sigma^2(M_h) = \int \Delta^2(k, z) |\tilde{W}(kR_f)|^2 d\ln k$ where $\tilde{W}(kR_f)$ is the Fourier transform of the filtering function. The Press-Schechter approach shows that the mass function has a power-law slope for the abundance of small haloes and an exponential cutoff for massive haloes. The Press-Schechter mass function provides a good approximation to direct results from N-body simulations.

The cosmological N-body simulation is the numerical method to solve the collisionless Boltzmann equation. This numerical method is theoretically well motivated as it seeks the solution of the Boltzmann equation by Monte-Carlo tracers (the method of characteristic). The phase space distribution function is represented by an ensemble of many 'particles';

each of them is a random realization of a point in phase space. The gravitational interaction (Poisson's equation) is solved by a fast solver such as Particle Mesh method, Tree method, or hybrid TreePM (e.g. [194, 28]). The initial condition is set usually by using the first-order Lagrangian perturbation theory (Zel'dovich approximation). The N-body method has become a primary theoretical tool because of its remarkable power to model structure formation in detail. For example, the Millennium simulation showed that the large-scale distribution of galaxies in the gravitational instability picture in ΛCDM cosmology explains the observed distribution very well ([422]).

Finally we remark on the difference in structure formation between low and high redshifts. In the CDM paradigm, structure formation proceeds hierarchically, i.e small scale nonlinear objects are formed before larger ones. The hierarchical build up is a result of the positive slope of the dimensionless power spectrum where the effective slope of power spectrum, $n_{\rm eff} = d\ln P(k)/d\ln k$, is in the range $-2 < n_{\rm eff} < 0$ at the scales of interest for galaxy formation at lower redshift. However, at higher redshift and smaller scales, the effective slope approaches asymptotically $n_{\rm eff} \to -3$ because of the turnover caused by the Mészáros effect. This means that a wide range of scales become nonlinear $\Delta^2(k) > 1$ nearly at the same time. This timing when Fourier modes enter nonlinear scales is also modulated by the large-scale perturbation, which causes the host haloes of first stars to form inside a progenitor of protocluster region at the present-day ([477, 13]). This $n_{\rm eff} \to -3$ limit suggests that the small scale early structure formation might have proceeded in a less hierarchical manner than at late time. The difference in the effective slope manifests the difference in the morphology of the large-scale structure between low and high redshifts ([166]). The mixing of a wide range of small scale Fourier modes could facilitate the rapid growth of early objects, as well as it makes the early large-scale structure to be more coherent than at late time ([166]).

1.5 Cosmological probes

The preceding § 1.3 and 1.4 reviewed the theoretical foundations of structure formation in ΛCDM cosmology. In this section I review the observational foundations of this paradigm and measurements of cosmological parameters.

1.5.1 H_0 measurement and the extragalactic distance scale

Since the influential works in 1920's (e.g. [201]), the direct measurement of the present-day expansion rate, H_0, has been the central interest to cosmological studies. The determination of H_0 was one of the key projects motivated to launch Hubble Space Telescope ([152] for review). The basic principle is to measure the recession velocity v of distant objects and the luminosity distance d_L independently, and then to determine H_0 from the relation $v = H_0 d_L$. While the recession velocity is readily available from the redshift measurement via spectroscopic observations, $v = cz$, the estimation of the distance must rely on a method for inferring the intrinsic luminosity L, hence the luminosity distance

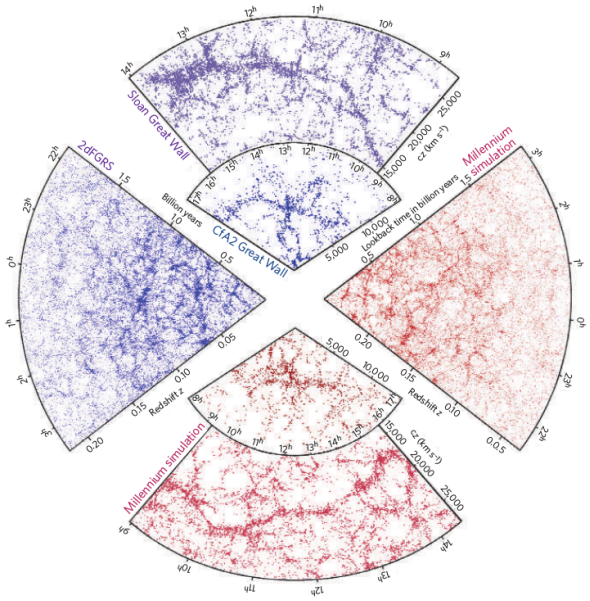

Figure 1.3: The observed (panels with blue points) and simulated (panels with red points) large scale structure in the galaxy distribution ([422]). The large-scale distribution of galaxies is remarkably similar between observation and simulation.

$d_L = \sqrt{L/4\pi F}$ where F is the observed flux. The primary objects used for this purpose are Cepheid variable stars and Supernova Ia (SNe Ia). Cepheids lie on the instability strip in Hertzsprung-Russel diagram, which exhibits the tight correlation between its pulsations period and the intrinsic luminosity ([478]). The peak luminosity of SNe Ia correlates with the rate it fades away ([152], see [192] for SNe Ia physics). Distance to Cepheids are first calibrated using main sequence fitting ([223]) or parallax ([24]). Cepheids, which are detectable by HST out to galaxies at $\lesssim 20$ Mpc, are used as a calibrator for SNe Ia, using nearby galaxies with detectable Chepheids and hosting SNe Ia. SNe Ia are then observed out to $\lesssim 400$ Mpc for the H_0 measurement.

Figure 1.4 shows the 2001 result from HST Key Project with the measured expansion rate of $H_0 = 71 \pm 2(\text{rand.}) \pm 6(\text{sys.})$ km s^{-1} Mpc^{-1} ([153]). Other distance indicators such as Tully-Fisher relation for spiral galaxies and Fundamental plane for elliptical galaxies were combined to improve the fidelity of the measured H_0 against astrophysical systematics related to any peculiarity of the objects ([153]). The H_0 measurement in 2011 reported a more precise value $H_0 = 73.8 \pm 2.4$ km s^{-1} Mpc^{-1} with $\sim 3\%$ uncertainty including their systematics error estimate ([372]). There are a wide range of sytematic uncertainties including Cepheid zero point calibration, metallicity, dust reddening, and peculiar velocities of host galaxies ([152, 176]), which could shift the central value of the measured H_0 larger than the quoted error budget (by $\sim 2.5\%$ in [133]). These direct measurement established the value of H_0 is ~ 70 km s^{-1} Mpc^{-1}, but a further improvement both in precision and accuracy would depend on the improved treatment and understanding of (astrophysical) systematics.

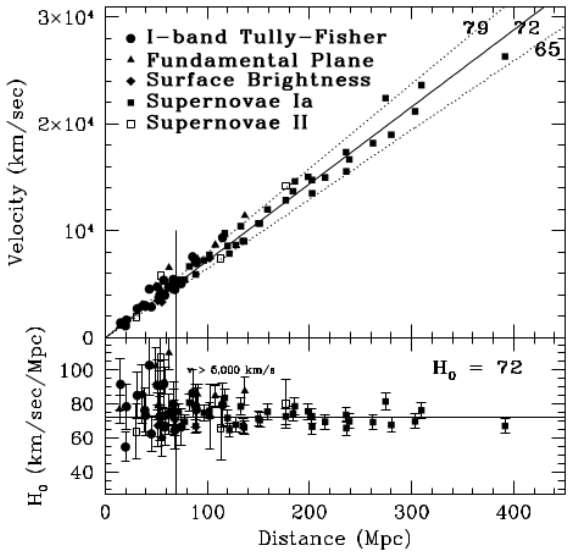

Figure 1.4: Hubble diagram of velocity vs distance for various distance indicators calibrated by Cepheids: SNe Ia, Tully-Fisher relation, Fundamental plane, surface brightness fluctuation of galaxies, and Type II supernova (Freedman et al. 2001).

1.5.2 Cosmic Microwave Background (CMB)

Since the initial accidental dicovery of CMB in 1964 by Penzias and Wilson ([344]), followed by the Cosmic Background Explorer (COBE) measurement of CMB blackbody temperature 2.735±0.06 K ([271]) and anisotropy ([412]), the hot big bang paradigm of CMB has established its place as a firm probe of cosmology (e.g. [200, 127, 232] for review). This is not only because accurate measurements from many dedicated intruments are available ranging from space-based missions such as Wilkinson Microwave Anisotropy Probe (WMAP) and Planck to ground-based telescopes such as Atacama Cosmology Telescope (ACT) and South Pole Telescope (SPT), but also because of the ability to theoretically model the CMB anisotropies cleanly from the physics of photon-baryon fluid in the early universe using the perturbed Boltzmann-Einstein equations ([258, 397, 251]). With the minimal flat ΛCDM model, the predicted CMB power spectrum involves only six-parameters, the dimensionless Hubble constant h, matter content $\Omega_m h^2$, baryon content $\Omega_b h^2$, the amiplitude of mass fluctuation σ_8, the primordial spectral index n_s, and Thomson optical depth τ. The excercise is to generate a family of cosmological model, then find the best-fit parameters to the observed CMB power spectrum.

CMB data alone has a tremedous power to constrain cosmological model; however, some internal degeneracies exist for the cosmological interpretation of the temperature and polarization power spectra. First, the effect of τ and σ_8 is to scale up and down the normalization of the temperature power spectrum. A larger optical depth due to electrons produced by reionization tends to dump the temperature anisotropy as the ionized medium is opaque to CMB photons. The change in the amplitude of primodal fluctuation σ_8 triv-

ially shifts both temperature and polarization power spectra. The $\tau - \sigma_8$ degeneracy can be broken with the polarization data, E-mode polarization power spectrum, because the Thomson scattering produces the polarized components in the CMB anisotropy. Second, the primary information on h, Ω_m, Ω_b comes from the acoustic peaks. While these parameters does not exhibit a strong degeneracy under the assumption of spatially flat geometry of the universe, relaxing the model to allow an open or closed (i.e $\Omega_{tot} < 1$ or > 1) universe produces the $\Omega_m - \Omega_\Lambda$ degeneracy (a.k.a. geometric degeneracy [49, 496]). External data on H_0 or Ω_m from SNe data or galaxy surveys are required to determine the geometry of the universe. Combining CMB with the H_0 measurement from SNe Ia and Cepheids ([153]), for example, lifts this degeneracy. Analysis of CMB with H_0 or galaxy survey data shows a strong preference to flat geometry, hence $\Omega_m \simeq 0.3$ and $\Omega_\Lambda \simeq 0.7$ model. CMB lensing, however, offers a way to constrain Ω_Λ internally within CMB data alone because the change in Ω_Λ, which affects the matter growth at later epoch (e.g. Einstein-de Sitter cosmology ($\Omega_\Lambda = 0$) has more structure than ΛCDM), propagates to a different gravitational lensing signal by the large-scale structure ([402, 350]).

A series of WMAP results ([418, 417, 233, 193]) has placed ever improving constraints on cosmological parameters. Figure 1.5 shows the temperature map of CMB and the measured and best-fit temperature power spectra for Planck mission, which constrains the cosmological parameters as $H_0 = 67.51 \pm 0.64$, $\Omega_\Lambda = 0.6879 \pm 0.0087$, $\Omega_m = 0.3121 \pm 0.0087$, $\sigma_8 = 0.8150 \pm 0.0087, \tau = 0.063 \pm 0.014$, and $n_s = 0.9653 \pm 0.0048$ using the temperature and E-mode polarization power spectra and CMB lensing ([348, 351]).

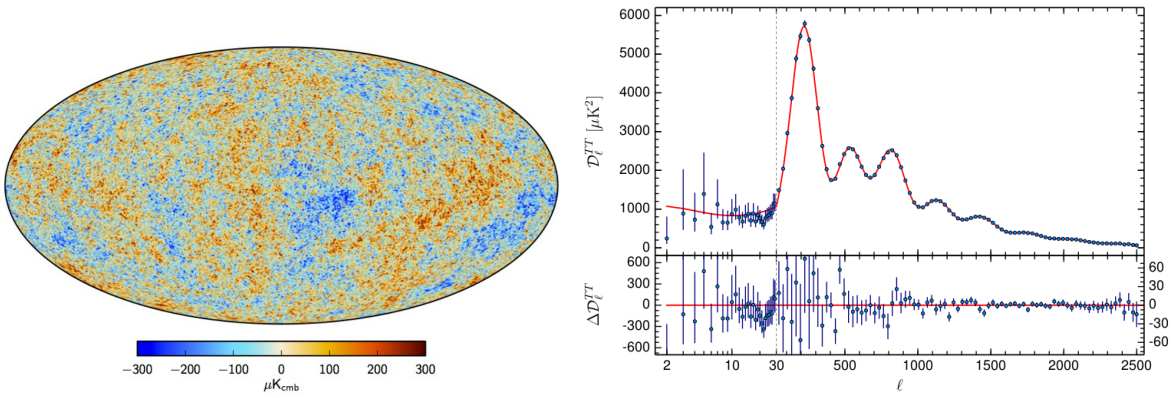

Figure 1.5: Left: The temperature map of the CMB. Right: The CMB temperature power spectrum (Planck collaboration 2015)

1.5.3 Large-Scale Structure: galaxy surveys and Lyα forests

The large-scale structure of the universe traced by the galaxy distribution and Lyα absorption lines in QSO spectra provide an observational window to study the cosmological model at redshift range $0 < z \lesssim 3$ (e.g. see [132, 345] for review). Figure 1.6[2] lists galaxy

[2] http://www.astro.ljmu.ac.uk/~ikb/research/galaxy-redshift-surveys.html

redshift surveys such as 2 degree Field Galaxy Redshift Survey (2dF; [85]), Sloan Degital Sky Survey (SDSS; [491]), and most recently Bayron Oscillation Spectroscopic Survey (BOSS; [105]) as a part of SDSS-III ([135]).

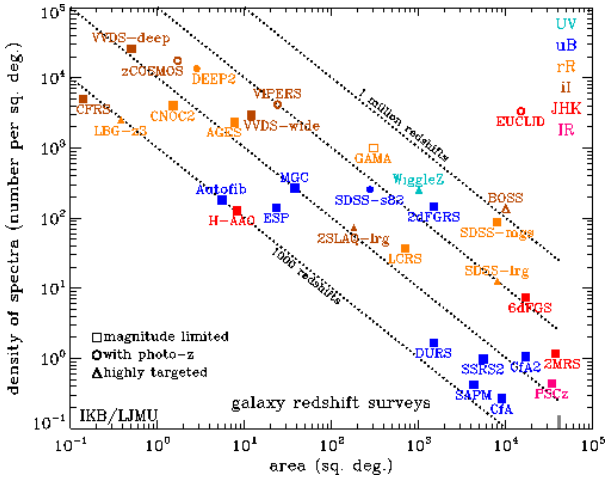

Figure 1.6: List of galaxy redshift surveys. (Credit: Ivan K. Baldry)

Galaxy redshift surveys probe the matter power spectrum at lower redshift. A constraint on the matter content $\Omega_m h$ comes from the turnover scale of the power spectrum at the horizon length at the matter-radiation equality (Meszaros effect), and the baryon fraction Ω_b/Ω_m can be constrained by the amplitude of baryon acoustic oscillation ([346, 84]). The normalization and very large-scale power contains the information about $b\sigma_8$ and n_s where b is galaxy bias parameter. By extending the analysis into full 2D redshift-space, the redshift-space distortion due to peculiar velocity of galaixes can constrain the growth rate of structure $\beta = f/b \simeq \Omega_m^{0.6}/b$ ([341, 188]). 2dF Survey has established that the matter conent of the universe is $\Omega_m \simeq 0.3$ through the shape of power spectrum and the redshift space distortion.

Since the discovery of BAO in galaxy redshift surveys ([84, 136]), more recent surveys (6dF; [30], WiggleZ; [34], BOSS; [8]) focus on the BAO signal in galaxy clustering. The acoustic scale measured from CMB experiment provide a standard ruler. The measurement of the acoustic scale in the transverse and line-of-sight directions directly constrains the angular diameter distance and Hubble parameter at various redshift ([473]). Anisotropic clustering in BOSS provided direct constraint on Hubble expansion rate at $z = 0.57$ as $H(z = 0.57) = 96.8 \pm 3.4$ km s^{-1} Mpc^{-1}, which is consistent with CMB data prediction for a flat ΛCDM cosmology ([8]).

In addition, Lyα forest in QSO spectra is a unique tool for probing the matter distribution at high redshift $z \sim 2 - 4$. Because the physics of the intergalactic gas is much more well understood than galaxy formation, the direct comparison between observations and cosmological hydrodynamical simulations of Lyα forest even in the 90's demonstrated the validity of the gravitational instability picture in the CDM paradigm and cosmic web of the large-scale structure, with effectively only one additional astrophysical parameter

for the UV background ([71, 190, 308, 500]). The measurement of 1D flux power spectrum from Lyα forest at $z \sim 2-4$ (Keck spectra; [96], SDSS; [275], BOSS; [330]) is consistent with $(\Omega_\Lambda, \Omega_m) \simeq (0.7, 0.3)$ model. As it provides the very different redshifts and scale to CMB and galaxy surveys, Lyα forest data provides a further support of ΛCDM cosmology. A recent series of analysis of BAO signal in 3D Lyα forest clustering (BOSS; [411, 61, 111]) extended the direct measurement of Hubble expansion rate to $z \simeq 2$ as $H(z = 2.34) = 222 \pm 7$ km s^{-1} Mpc^{-1} ([111]). Interestingly, BOSS DR11 Lyα forest data exhibits $\approx 2.5\sigma$ discrepancy with the best-fit Planck parameters ([111]). Whether it is caused by systematics due to continuum fitting, inaccurate astrophysical modelling, or true deviation of cosmological origin is still unclear. With available data at hand, ΛCDM model and general relativity remains a good description of cosmology and the theory of gravity responsible for the large-scale structure in the universe.

Overall, based on the current data, analysis of H_0 measurement, CMB, and redshift surveys of galaxies and Lyα forests shows that ΛCDM cosmology is highly consistent with observations. Thus, astrophysics of galaxies and the IGM can be posed as an initial condition problem with the well constrained parameters in ΛCDM cosmological framework.

1.6 Astrophysics of Galaxies and the IGM

The begining of extragalactic astronomy dates back to the time when the realm of fuzzy nebula on the sky are discussed about their origin; whether they are collection of billions of stars located outside our Milky Way galaxy or not (Shapley-Curtis Debate in 1920). A historical overview is useful to highlight the paradigm shift in understanding the astrophysics of galaxies and the IGM ([137]). This is not a rigrous chronology, but concentrates on the gradual shifts in the focus of community.

Before 1980, the classic view of galaxies as isolated islands in the universe was held. Galaxies are thought to form through the evolution and formation of stars and gas inside galaxies in isolation. The understanding was based on stellar population synthesis modelling built on the well-matured field of stellar evolution ([446] for review).

Around 1980-1995, the basis of the current paradigm is set – hierarchical galaxy formation in ΛCDM cosmology – the stellar content of galaxies is thought to be built in the bottom-up manner by cooling and condensation of gas inside dark matter haloes ([476]), whose dynamics and formation is determined by the gravitational instability during the large-scale structure formation in the universe (§ 1.4). This shifted the isolated island paradigm to a dynamic interacting picture of galaxy formation through mergers, which transform the galaxy morphological type and build up stellar mass.

Recent progress since 1995 has placed increasing emphasis on the role of feedback and the environment in shaping galaxies. Theory has been pushed for self-consistent treatment of feedback and baryonic processes of galaxies in a cosmological hydrodynmical framework ([424, 196, 389, 469]). The paradigm centers on the regulation of star formation inside galaxies by stellar and AGN feedback and the gas acretion from the intergalactic environment to feed subsequent star formation. Both the host halo mass and the galaxy

environment (field, group, and cluster regions) are thought to be resonsible for various star formation quenching mechanisms, which controls the morphology-density relation as well as galaxy downsizing. Furthermore, the high-redshift frontier for first galaxies and reionization has been undergone the period of renaissance, which is driven by a surge of observational surveys targetting the high-redshift universe.

1.6.1 Physical processes in galaxies

In this section I review the physical processes responsible for shaping the properties of galaxies. For a more detailed account, I refer readers to many exellent reviews and books that summerize the physical processes in galaxies in a cosmological framework (e.g. [311, 414, 25, 408]). The basic physical ingredients for shaping the properties of galaxies (the cooling of gas, star formation and feedback, and black hole formation and feedback) are relavant for both low-redshift galaxies and early galaxies during reionization.

Cooling of gas and gas accretion

For the gas inside dark matter haloes to form stars, it must first cool to the temperature at which the gravitational force exceeds the thermal pressure support. When the primodial gas falls into haloes, the gas is shock heated to the virial temperature, $3k_B T_{\rm vir}/(2\mu m_p) \sim GM_h/r_{\rm vir}$[3] $T_{vir} \sim 2.8\times10^6(1+z)(M_h/10^{12}\ {\rm M}_\odot)^{2/3}(\mu/0.6)$ K, where μ is the mean molecular weight. A criterion for whether a gas cloud can cool efficiently and keep collapsing is given by the ratio between the cooling timescale t_{cool} and the free-fall timescale $t_{ff} = \sqrt{3\pi/(32G\rho)}$. When the cooling timescale is shorter than the free-fall timescale, a gas cloud cools fast enough for runnaway gravitational collapse. On the other hand, when $t_{cool} > t_{ff}$, the gas forms a quasi-hydrostatic atmosphere of hot gas.

The various radiative processes control the cooling of gas. At $T_{vir} > 10^7$ K, the main dissipation mechanism is bremsstrahlung (scattering of free electrons by atomic nuclei). At $10^4 < T_{vir} < 10^7$ K, collisional ionization and excitation of atoms is a main cooling channel (atomic cooling). Below $T_{vir} < 10^4$ K gas is mainly neutral and can only cool through molecular processes (molecular cooling) or metal line cooling. The cooling rate at $T_{vir} \lesssim 10^7$ K depends strongly on the metallicity of the gas, which could differ by an order of magnitude between primordial and solar composition ([25]). A large metallicity allows the gas to cool much more efficiently than primordial gas.

Different cooling mechanisms are responsible for the formation of different galaxies at various mass and redshift. For a galaxy at $z = 2$ with Milky-Way size halo 10^{12} ${\rm M}_\odot$, the main cooling channels are the atomic processes in the metal enriched gas. The first galaxies at $z \sim 10-15$ with halo mass 10^8 ${\rm M}_\odot$ cool largely via hydrogen atomic cooling. First stars, which are key progenitors for shaping first galaxies, are formed in minihaloes $M_h \sim 10^6$ ${\rm M}_\odot$ through molecular cooling via H_2 ([58]).

[3]Here, $r_{\rm vir}$ is in the physical unit, $r_{\rm vir} = R_{\rm vir}/(1+z)$.

There are two modes of gas accretion onto galaxies. Firstly, when $t_{cool} > t_{ff}$, a quasi-hydrostatic atmosphere needs to cool at its centre[4], and the gas accretes onto the central star-forming region. This mechanism is called hot mode accretion. On the other hand, when the cooling timescale is shorter than the free fall timescale, the intergalactic gas can penetrate into a central galaxy without ever been shock heated ([475, 33]). This form of gas flow occurs along the filamentary structure of cosmic web and is referred to as cold stream (or cold mode accretion) ([224, 110, 461]). In addition to the above two smooth modes of gas accretion, major mergers can trigger more violent feeding of the gas into central kpc regions, resulting in starbursts ([51]). These modes of gas accretion are important for the subject of first galaxies and reionization because a simple extrapolation of lower redshift studies suggests that the stellar contents of early galaxies driving reionization could also be assembled by cold stream-fed star formation.

Star formation and feedback

As gas accretes onto the central regions of the haloes (a site of galaxy formation), the self-gravity of gas starts to dominate over that of dark matter. The collapse and fragmentation of gas form molecular clouds and multi-phase interstellar medium (ISM). Giant molecular clouds span mass scale of $10^5 - 10^6$ $M_\odot$ and length of $10 - 100$ pc. Formation of star clusters takes place in the clumps in molecular clouds where the physics of turbulence, magnetic fields, and self-gravity are thought to be key players for star formation (e.g. [279, 238] for review). Stars typically form as binary or multiples. While stellar evolution of single isolated stars is well understood ([231] for book), if the seperation between two stars is small enough to interact through stellar winds or Roche lobe overflow, the course of stellar evolution is also altered. While star formation theory (the rate of star formation in individual molecular clouds) and stellar astrophysics (the evolution and death of single and binary stars) are themselves active fields of research, galaxy formation theory focus on the role of large-scale properties (e.g. stellar mass, metallicity, gas accretion rate, merger) on the rate and history of star formation in a galaxy as a whole.

There are two modes of star formation in galaxies: star formation main sequence and starburst. Star formation main sequence refers to a tight correlation between SFR and M_*, which is driven by the steady formation of stars in disk-like galaxies. Starburst is a sudden rapid formation of stars which may be triggered by mergers or any other instability mechanism (e.g. [408] for review).

Stellar feedback impacts the gas around young star clusters in a variety of ways. Massive stars play an pivotal role in feedback during the entire evolutionary stages. At early stages, they feedback to the surrounding medium through protostellar jets, ionizing radiations, stellar winds and radiation pressure ([198]). At a late stage, the death of massive stars as supernova (SNe) influences the ISM. Ejecta from supernova sweep the ISM away through shock and quickly convert the kinetic energy into thermal energy. The form of

[4]Because the main cooling channels of hot gas are two-body processes, the cooling is faster at high density regions as it scales as $\propto n_{gas}^2$.

feedback can be classified into mecanical, thermal, radiative, and chemical feedback. Mechanical feedback is driven by momentum injection to the ISM. Thermal feedback inject energy, rising the temperature. Radiative feedback is associated with the ionization of atoms and the dissociation of molecules in the ISM. Chemical feedback inject metals from stellar nucleosynthesis into the ISM. Note that there is no border line between the form of feedback. For example, photoionization by UV radiation from massive stars creates H II regions (radiative feedback). At the same time, the photoionization heating increases the temperature to $\sim 2 \times 10^4$ K (thermal feedback). Supernova explosion triggers both thermal, and mechanical feedback as well as enriching the ISM by metals. Hydrodynamical zoom-in simulations resolving the scale of molecular clouds theoretically support this paradigm, where the multitude of feedback processes nonlinearly interacts and is responsible for shaping the star formation and stellar mass assembly of galaxies ([196]).

An insightful concept for star formation in galaxies is self-regulation. The local star formation rate in dense molecular clouds is $\dot{\rho}_* = \epsilon_* \rho_{gas}/t_{ff} \propto \rho_{gas}^{3/2}$, where star formation efficiency is $\epsilon_* \sim 1\%$ ([239]). The timescale of giant molecular clouds (GMC) to form young star clusters is $\sim 10^7$ yr, which is much shorter than the age of galaxies $\sim 10^9$ yr ([311]). This means that as soon as molecular clouds are formed inside galaxies, they immediately form stars, and the rate at which GMC is formed is therefore a limiting step controlling the rate of star formation. Stellar feedback from massive stars dispels the GMC as well as the Lyman-Werner radiation from massive stars disscoiates molecular hydrogen. This limits the amount of gas available for further star formation. Because the bottleneck process is controlled by feedback, we can tink of star formation in galaxies as a system self-regulated through stellar feedback ('feedback-regulated star formation') ([197]). This causes the galaxy-wide SFR to be insensitive to small-scale efficiency to convert molecular clouds to stars, but sensitive to the feedback mechanism of massive stars ([197]). We can push this picture of self-regulation further by placing it in a cosmlogical context where the intergalactic gas accretion onto galaxies from the large-scale structure. Because the internal timescales of the ISM of galaxies to form molecular clouds and stars are shorter than Hubble timescale, fuelling of gas from the intergalactic environment is critial to sustain the galaxy-wide star formation over cosmic timescales. Therefore, the gas accretion from the IGM and the feedback-regulation of star formation are two key processes for controlling the star formation history of galaxies ([390]).

Black hole formation and growth and AGN feedback

Supermassive black holes (SMBHs) of mass $\sim 10^6 - 10^9$ $M_\odot$ are ubiquitous in the centre of galaxies. In a unified model of Active Galactic Necluei (AGN), a SMBH powers the accretion disk, which emits a large amount of radiation by the non-thermal process outshining the stellar radiation from the host galaxies during its activity.

The flow chart presented by Martin Rees in 1978 ([368]) succinctly describes the possible pathways leading to the formation of SMBHs ([471] for recent review). The onset of SMBH formation is seeded by (i) the death of massive stars as the remnants of Pop III stars, (ii) compact star clusters, and/or (iii) direct collapse of a gas cloud. These seed black holes

subsequently grow by gas accretion and mergers with other black holes.

The AGN feedback influences the star formation in galaxies in various ways. For example, the UV and X-ray radiation emitted from accretion disk around a SMBH and hot corona, as well as Compton scattering of high-energy photons by ionized gas ([83]), heat up the ISM, which prevents further formation of stars. The radiation from AGN may also exert radiative force to drive the winds. The AGN feedback is a leading explanation for the exponential break in the galaxy luminosity function and the quenching of star formation in massive galaxies ([408]). These insights from lower redshift studies are viable for understanding the formation of first quasars and its relation to host galaxies at $z > 6$. AGNs (and mini-quasars) may play a important role during reionization through their UV and X-ray radiations. The reionization in QSO environment is a topic discussed in detail in Chapter 3.

1.6.2 Physical processes in the IGM

In the local universe, galaxies are only a minor $\sim 7\%$ reservoir of the total baryons in the universe ([155, 405]). The rest $\sim 90\%$ of baryons resides in the space between galaxies in the form of the intergalactic medium (IGM). The evolution of the IGM is closely linked with the formation of galaxies through gas accretion and feedback/galactic winds. The sub-region of the IGM around galaxies, called circum-galactic medium (CGM), is the interface between the IGM and galaxies, where the baryon cycle regulates the evolution of galaxies. On larger scale, galaxies and AGNs influence the IGM through cosmic reionization. In a broad sense, the reionization of the IGM influences the entire cosmic history from as high as $z \sim 15-20$ through the first reionization of hydrogen and helium to $z \sim 2-4$ through the second reionization of helium. The photoheating of the gas through the reionization process raises the cosmological Jeans mass, and therefore suppresses the formation of small dwarf galaxies. The fluctuations in the UV background and temeperature in the IGM driven by the reionization processes can act as a source of astrophysical systematics to study cosmology using the Lyα forest at $z \sim 2-4$ ([244, 283, 352, 178]). The astrophysics of the IGM thus occupies a unique important place in the study of galaxy formation, reionization, and cosmology.

The density fluctuaitons of the IGM trace closely the matter distribution formed through the CDM-dominated gravitational field, $\rho_b = (\Omega_b/\Omega_m)\bar{\rho}(1+\delta)$. This is because the hydrodynamic pressure only acts on scales smaller than the cosmological Jeans length $\lambda_J = c_s \pi^{1/2} (G\rho_m)^{-1/2}$ where $c_s \simeq \sqrt{kT/m_p}$ is the sound speed and ρ_m is the physical matter density,

$$\lambda_J \approx 110 \Delta_m^{-1/2} \left(\frac{\Omega_m h^2}{0.142}\right)^{-1/2} \left(\frac{T}{10^4\ \mathrm{K}}\right)^{1/2} \left(\frac{1+z}{5}\right)^{-3/2} \ \mathrm{pkpc}. \tag{1.12}$$

The ionization state of hydrogen in the IGM at the intermediate redshift $2 < z < 5$ is determined by the photo-ionization equilibirum with the metagalactic UV background (the angle-averaged specific intensity J_ν) built from star forming galaxies and QSOs, The

H I-photoionization rate $\Gamma = 4\pi \int \sigma_{\rm HI} J_\nu/(h\nu) d\nu$ and the case-A recombination rate of hydrogen $\alpha_A(T) = 4.063 \times 10^{-13} (T/10^4{\rm K})^{-0.72} {\rm cm}^3\ {\rm s}^{-1}$ establish the photoinization equilibrium $\Gamma n_{\rm HI} = \alpha(T) n_e n_{\rm H\,II}$ where $n_{\rm H} = \rho_{\rm crit}(0)\Delta_b(\Omega_b/m_p)(1-Y)(1+z)^3$ and the $Y \simeq 0.25$ is the primordial helium fraction by mass. The neutral fraction of hydrogen is then given by

$$x_{\rm HI} \simeq \frac{\alpha_A(T) n_{\rm H} f_e}{\Gamma} \approx 9.5 \times 10^{-6} \Delta_b f_e \left(\frac{T}{10^4\ {\rm K}}\right)^{-0.72} \left(\frac{\Gamma}{10^{-12}\ {\rm s}^{-1}}\right)^{-1} \left(\frac{\Omega_b h^2}{0.022}\right) \left(\frac{1+z}{5}\right)^3, \tag{1.13}$$

where f_e is the fraction of electron per hydrogen atom, which is $f_e = \frac{2Y}{4(1-Y)}\ \left[\frac{Y}{4(1-Y)}\right] \approx 1.174$ [1.087] after [before] He II reionization. The dependence on redshift comes from the cosmic expansion. Hydrogen in the IGM is highly ionized at $2 < z < 5$ by the metagalactic UV background, but residual neutral hydrogen $x_{\rm HI} \sim 10^{-5}$ remains.

The temperature of the IGM is mainly controlled by photoionization heating, adiabatic cooling, and the Compton cooling of the gas with the CMB. This thermodynamics of the IGM can be compared to the cooling for galaxy formation where neither Compton cooling nor adiabatic cooling is dominant. For gas inside haloes, the high temperature by shock heating and the high density allow cooling through the two-body (atom-atom, atom-electron) collisional processes. However, because the IGM has a primordal chemical composition and low density, one-body processes (atom-photon) such as compton cooling and photoionization are only available mechanisms for cooling and heating. The evolution of the temperature for a Lagrangian fluid element is (e.g. [282] for review)

$$\frac{dT}{dt} = -2HT + \frac{2T}{3\Delta_b}\frac{d\Delta_b}{dt} + \frac{2(\mathcal{H}-\mathcal{C})}{3k_B n_b}, \tag{1.14}$$

where $\mathcal{H}$ ($\mathcal{C}$) is the heating (cooling) rate per volume (in units of erg ${\rm s}^{-1}$ ${\rm cm}^{-3}$). On the right hand side, the first term is adiabatic cooling by cosmological expansion and the second term is adiabatic heating/cooling by the local compression/expansion of a Lagrangian element. The temperature quickly establishes the tight power-law relation between temperature and density $T = T_0 \Delta_b^{\gamma-1}$ due to the combination of photoionization, Compton cooling and adiabatic cooling ([310, 203, 441]). The photoionization heating dumps the amount of energy, $\langle E_{\rm HI}\rangle = \int_{\nu_L} (h\nu - h\nu_L)\frac{4\pi I_\nu}{h\nu}\sigma_{\rm HI} d\nu \Big/ \int_{\nu_L} \frac{4\pi I_\nu}{h\nu}\sigma_{\rm HI} d\nu \propto h\nu_L$, per photoionization into the gas. The heating rate by photoionization is $\mathcal{H} = \langle E_{\rm HI}\rangle \Gamma n_{\rm HI}$. As the neutral fraction of hydrogen is set by the photoionization equilibrium, $n_{\rm HI} \approx \alpha_A(T) n_{\rm H}^2/\Gamma$, the heating of the IGM by the photoionization of residual neutral hydrogen is $dT/dt = 2\mathcal{H}/(3k_B n_b) \approx 2\langle E_{\rm HI}\rangle \alpha_A(T) n_{\rm H}/3k_B$. The slope of the asymptotic temperature-density relation is mainly determined by the balance between the cosmological adibatic cooling and photoionization heating of residual neutral hydrogen, $2HT = 2\langle E_{\rm HI}\rangle \alpha_A(T) n_{\rm H}/3k_B$, which implies that $T \propto T^{-0.72}\Delta_b$. The asymptotic slope of temperature-density relation is then

$$T = T_0 \Delta_b^{\gamma-1} \propto \Delta_b^{1/1.72}. \tag{1.15}$$

This simple estimate of the asymptotic slope $\gamma = 1/1.72 + 1 = 1.58$ is consistent with the observed range and hydrodynamical simulations ([39, 38]). Strictly speaking, this

derivation is incomplete, because it ignored the adiabatic heating/cooling by local gas element. During the structure formation, lower density region can preferentially cooled through adiabatic process $T \propto \Delta_b^{2/3}$. This in practice contributes to form the asymptotic slope (fortunately, the result of a more detailed derivation is nearly identical to the value derived above (for more detail discussion, see [288]). While the asymptotic slope is determined by the adiabatic cooling and photoionization heating, the Compton cooling plays an important role for the relaxation speed for approaching the asymptotic $T - \Delta_b$ relation. The timescale for Compton cooling is $t_{\rm comp} = 3m_e c/[4\sigma_T a_R T_{\rm CMB}(1+z)^4]$. Compared to the adiabatic cooling timescale by cosmic expansion $t_{\rm ad} = [2H(z)]^{-1}$, the ratio is $t_{\rm comp}/t_{\rm ad} \approx 1.3(\Omega_m h^2/0.142)^{1/2}[(1+z)/7]^{-5/2}$. The Compton cooling is comparable to adiabatic cooling at $z = 6$. The Compton scattering and cosmological expansion together quickly cool the IGM and erase the memory of photoionization heating by H I reionization, i.e. the first passage of H I ionization fronts, when the slope was nearly isothermal $\gamma \approx 1$ with dispersion. The IGM temperature approaches towards the tight asymptotic $T - \Delta_b$ relation.

This photoionization heating by the UV background causes Jeans pressure smoothing in the IGM. Before the photoionization heating by reionization, the temperature of the IGM is low $T \sim 10$ K ([396]) in absence of pre-heating. Therefore the Jeans length is as low as ~ 3.4 pkpc at $z = 10$. This means that the IGM is more clumpy before the passage of ionization fronts. The photoionization heating helps to keep the universe reionized by reducting the clumping factor ([338]).

When the density fluctuation is high enough to be opaque against ionizing photons, the gas clumps can be self-shielded against the UV background. The column density of absorber, at which gas starts to be self-shielded is when the optical depth of the absorber reaches a unity, $\tau = \sigma_{\rm HI} N_{\rm HI} > 1$, where $\sigma_{\rm HI} = 6.3 \times 10^{-18}$ cm^2, is $N_{\rm HI} \gtrsim 1.6 \times 10^{17}$ cm^{-2}. A characteristic size of photoionized Lyα absorption systems is of order Jeans length (Jeans argument, [388]). Thus, the photoionized Lyα forest absorbers have H I column density $N_{\rm HI} = x_{\rm HI} n_{\rm H} \lambda_J$ where $x_{\rm HI}$ is given by photoionization equilibrium,

$$N_{\rm HI} \approx 7.5 \times 10^{13}\ {\rm cm}^{-2} f_e \Delta_b^{3/2} \left(\frac{T}{10^4\ {\rm K}}\right)^{-0.22} \left(\frac{\Gamma}{10^{-12}\ {\rm s}^{-1}}\right)^{-1} \left(\frac{\Omega_b h^2}{0.022}\right) \left(\frac{1+z}{5}\right)^{9/2}. \quad (1.16)$$

The critical overdensity at which the gas starts self-shielding $N_{\rm HI} > 1.6 \times 10^{17}$ cm^{-2} is ([158])

$$\Delta_{\rm ss} \approx 166 f_e^{-2/3} \left(\frac{T}{10^4\ {\rm K}}\right)^{0.146} \left(\frac{\Gamma}{10^{-12}\ {\rm s}^{-1}}\right)^{2/3} \left(\frac{1+z}{5}\right)^{-3}. \quad (1.17)$$

At $z = 4$, the outskirt of haloes of overdensity $\Delta \sim 150$ and the filamentary structure of the cosmic web make up the self-shielded gas in the IGM. The properties of the absorbers are conventionally classified into three classes: Lyα forest absorbers are diffuse photoionized IGM with $N_{\rm HI}/{\rm cm}^{-2} < 10^{17}$, the Lyman limit systems are the transition component between photoionized gas and fully self-shielded gas with $10^{17} < N_{\rm HI}/{\rm cm}^{-2} < 10^{20.3}$, and damped Ly$\alpha$ system with $N_{\rm HI}/{\rm cm}^{-2} > 10^{20.3}$ is thought to raise from clumps in galaxies or massive disk.

1.6.3 Radiation Hydrodynamics of Galaxies and the IGM

The above sections have reviewed the physical processes shaping galaxies (§ 1.6.1) and the IGM (§ 1.6.2) on the heuristic ground. The full treatment of the co-evolution of galaxies and the IGM in principle involves the well-known physics of gravity, hydrodynamics, and radiative transfer. The quesion is to understand how such complex interplays between galaxies and the IGM emerge from seemingly simple laws of classic physics in a cosmological framework.

The equations of radiation hydrodynamics

Firstly, we introduce the fundamental laws of radiation hydrodynamics. The equations of radiation hydrodynamics are described by the set of conservation laws. The hydrodynamics, i.e. the density ρ and velocity field $\boldsymbol{v}$ of the gas, is given by the Euler equations (e.g. [421, 440]),

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}) = 0, \tag{1.18}$$

$$\frac{\partial (\rho \boldsymbol{v})}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{v}\boldsymbol{v}) = -\boldsymbol{\nabla} P - \rho \boldsymbol{\nabla} \phi, \tag{1.19}$$

$$\frac{\partial (\rho e)}{\partial t} + \boldsymbol{\nabla} \cdot [(\rho e + P)\boldsymbol{v}] = -\rho \boldsymbol{v} \cdot \boldsymbol{\nabla} \phi + \mathcal{H} - \mathcal{C}, \tag{1.20}$$

where e is the total energy per unit mass, which includes the contributions from the internal energy per unit mass u (specific internal energy) and the kinetic energy of bulk flow, $\rho e = \rho u + (1/2)\rho v^2$. The equation of state for an ideal gas is needed to close the system. The equation of state relates the thermal pressure with a density and a specific internal energy, $P = (\gamma - 1)\rho u$, where γ is the adiabatic index ($\gamma = 5/3$ for a monatomic gas). The temperature of gas is readily given by $P = \rho k_B T / \mu m_p$. The gravity is the most important force in astrophysical systems. The gravitational potential ϕ is sourced both by gas and dark matter fluctuations in the large-scale structure of the universe through the Poisson's equation (§ 1.4). The heating and cooling functions, $\mathcal{H}$ and $\mathcal{C}$, couple the hydrodynamics to the various radiative processes, where the physics of radiative transfer and chemistry play a role.

The radiation field (specific intensity I_ν) is a macroscopic description of photons with a various range of electromagnetic spectrum at energy $h\nu$. The equation of radiative transfer is just a different way of expressing the Boltzmann equation of photons,

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \boldsymbol{n} \cdot \boldsymbol{\nabla} I_\nu = -\alpha_\nu I_\nu + \varepsilon_\nu, \tag{1.21}$$

where α_ν is the opacity and ε_ν is the emissivity. $\boldsymbol{n}$ is the direction of photon propagation. The first two angular moments of the specific intensity correspond to the angle-averaged intensity, $J_\nu = (1/4\pi) \int I_\nu(\boldsymbol{n}, \nu) d^2\boldsymbol{n}$, and the radiation flux, $\boldsymbol{F}_\nu = \int \boldsymbol{n} I_\nu(\boldsymbol{n}, \nu) d^2\boldsymbol{n}$. The photoionization heating of hydrogen injects the thermal energy to the gas by,

$$\mathcal{H}_{\rm HI} = n_{\rm HI} \int_{\nu_L}^{\infty} (h\nu - h\nu_L) \frac{4\pi J_\nu}{h\nu} \sigma_\nu d\nu. \tag{1.22}$$

Because the photoionization of He I and He II also heat up the gas, the total heating rate is $\mathcal{H} = \mathcal{H}_{\rm HI} + \mathcal{H}_{\rm HeI} + \mathcal{H}_{\rm HeII}$. The major cooling processes for hydrogen and helium are collisional excitation/ionization cooling and recombination cooling, which takes the form $\mathcal{C}_i = n_e n_i \beta_i(T)$ where the cooing rate coefficients $\beta_i(T)$ are tabulated in [69, 441, 447]. At a higher temperature, the gas predominantly cools through the Bremstrahlung emission $\mathcal{C}_{\rm ff} \approx 1.42 \times 10^{-27} T^{1/2} n_e (n_{\rm H\,II} + n_{\rm He\,II} + 4 n_{\rm He\,III})$ erg s^{-1} cm^{-3} (e.g. [441]). Compton cooling by scattering of free electrons by the CMB photons is important for high redshift and low density gas $\mathcal{C}_{\rm comp} = [4\sigma_T a_r k_B/(m_e c)] n_e T_{\rm CMB}^4(z)[T - T_{\rm CMB}(z)]$ ([295]). Molecular cooling and metal line cooling are more relavant on galactic scales (§ 1.6.1).

The ionization state of the gas must also be understood to determine the heating and cooling rate as well as to describe the reionization process. The rate equations involve the non-equilibrium chemistry,

$$\frac{\partial n_i}{\partial t} + \boldsymbol{\nabla} \cdot (n_i \boldsymbol{v}) = \alpha_{i,j}(T) n_e n_j - n_i \Gamma_i, \quad i = 1, \ldots, N_s \tag{1.23}$$

where the index i runs for all species, $\alpha_{i,j}(T)$ are the recombination cofficients, and the photoionization rate of the chemical element i is given by

$$\Gamma_i = \int_{\nu_L}^{\infty} \sigma_i(\nu) \frac{4\pi J_\nu}{h\nu} d\nu. \tag{1.24}$$

The radiation hydrodynamics of astrophysical fluid flows is fully specified by the Euler equations (1.18),(1.19),(1.20), radiative transfer equation (1.21), and the rate equations for non-equilibrium chemistry (1.23), together with the equation of state. In practice, galaxy formation simulations must be supplemented by a recipe to convert gas into stars/black holes and their radiation to source the emissivity in the radiative transfer equation. Once specified, radiation hydrodynamics of astrophysical fluid flows can be studied in detail on the theoretical ground. Because of the high complexity of the problem, we need to resort to nuermcal techniques and various approximations to solve the equations of radiation hydrodynamics. It is however important to keep always in mind the 'out-of-box' way of thinking; simulation is only 'a' tool for theorists.

Numerical methods

Solving the full problem of radiation hydrodynamics of galaxies and the IGM is very demanding. Various numerical algorithms and physical approximations must therefore be employed.

In numerical hydrodynamics, there are two widely used classes of algorithms; smoothed particle hydrodynamics (SPH, [420, 472]) and grid-based hydrodynamics with adaptive mesh refinement (AMR, [237, 60, 154, 439]). SPH take the Lagrangian viewpoint, which has an advantage in its automatic spatial adaptivity and ease to achieve high dynamic range. AMR takes the Eulerian viewpoint, which has arguably mathematically rigorous treatment and an ability to handle discontinuities well. Recent advances shrink these traditional numerical differences in hydrodynamic solvers between SPH and AMR methods.

Radiative transfer also employs a range of numerical algorithms. The methods that directly solve the radiative transfer equation (1.21) are the ray tracing algorithms, which can be classified into long ([2, 481]) and short characteristic ([297, 336]) methods as well as Monte Carlo method ([270]). All the methods attempt to realize the radiation field as a bundle of rays. As these methods involve the discretization in space and frequency as well as angular direction, the computational cost is usually very expensive. With a limited computational resource, they are typically employed for post-processing calculation[5] and to investigate the radiative transfer physics in detail. An alternative method is the moment method ([174, 149, 370, 376]), which reduces the dimensionality of the problem by taking the angular moments of the radiation field. This is similar to the fluid approach for the Boltzmann equation; however, it produces an infinite hierarchy of moment equations unless the system is closed by supplementing the higher order moment, e.g. radiation pressure tensor, in terms of the lower moments, e.g. radiation energy and flux, or by a direct solution of the ray equation (1.21). Despite these downsides, the moment method is typically computationally cheaper than the ray tracing methods. They are often used for full radiation hydrodynamical calculations.

The physical approximations widely used for cosmological hydrodynamical simulations are the optically thin limit approximation for the UV background radiation and the ionization equilibrium. When hydrogen and helium are already photoionized they are optical thin to the UV background, which eliminates the need to solve the radiative transfer equation. One then provides the homogenous UV background as a function of redshift, $J_\nu(\nu, z)$, which is modelled based on observations of the ionizing sources (galaxies and AGNs) and the sinks (Lyman-limit systems and damped Lyα systems) ([184, 142]). The ionization equilibrium assumption eliminates the need to solve the non-equilibrium chemistry. The ionization state of the gas given a metallicity readily determines the cooling function pretabulated in e.g. [432]. Recent cosmological hydrodynamical simulations employ these approximations ([389, 470]).

Another approximation is to ignore the radiative feedback on gas dynamics through photoionization heating. In this limiting regime, the radiative transfer and the ionization calculations (1.21) and (1.23) can be decoupled from hydrodynamical calculations (1.18), (1.19), (1.20), and can be performed in a post-processing manner. This approach (hydrodynamics + post-processing radiative transfer) is widely used in studies for IGM reionization (e.g. [81, 17]) and the escape fraction of ionizing radiation from galaxies (e.g. [259]).

Full radiation hydrodynamical simulations of cosmological reionization have now become possible ([173, 337, 317]). This is very exciting progess in computational astrophysics. It is worth to emphasize that all three methodologies, (1) hydrodynamical simulation, (2) hydrodynamics + post processing radiative transfer, (3) full radiation hydrodynamical simulation are complementary in understanding extragalactic phenomena involving galaxies and the IGM. Given the limited computational technology available, hydrodynamical calculations can create a 'suite' of simulations with various star formation and feedback

[5]Recently, the short characteristic method in SPH ([337]) is successfully applied for a full radiation hydrodynamical simulation of reionization.

physics with high dynamic range. The hydrodynamics+post-processing radiative transfer approach can now also be routinely performed to produce a radiative transfer simulation suite, where we can extensively study the multi-frequency radiative transfer physics and the role of galaxies and AGNs in the reionization process with various source models. Full radiation hydrodynamical simulations, of course, enable us to study the coupling of hydrodynamical and radiative transfer processes in galaxy formation and the IGM in full detail.

1.7 Bridging theory and observations

Stellar population synthesis, Lyα forest, Lyα transfer, and 21cm

There are a range of the important techniques to bridge between theory/simulation and observations. The stellar population synthesis (e.g. [385] for review) is an essential tool to connect the direct observable properties, i.e. spectrum and broad-band magnitudes, to the physical properties of galaxies directly given by radiation hydrodynamical simulations.

Synthetic Lyα forest spectra from simulations is an important direct tracer of the physical state of the IGM.

Monte-Carlo Lyα transfer modelling is now central in interpreting the Lyα emitting galaxies and the diffuse haloes, which contains information about the star formation, galactic winds, CGM, and reionization.

21cm signal is, once detected, the most direct observable to probe the state of the neutral IGM through its power spectrum and tomography. Understanding the physics of these tracers is as important as the radiation hydrodynamical simulations of galaxies and the IGM. They together form a vital theoretical pillar for interpreting observations of the high-redshift universe.

1.8 Observations of the High-Redshift Universe

The theoretical picture for astrophysics of galaxies and the IGM described in § 1.6 must be continously confronted with observations to build up a true picture of the physical world. I briefly review the observational insights of the high-redshift universe, with focus on the high-redshift galaxy properties and the IGM relavant to understand the Epoch of Reionization.

1.8.1 General demographics of galaxies

First of all, I briefly look into the redshift evolution of the galaxy population in general. Observational surveys of galaxies, when analysed with population synthesis modelling, show that the specific star formation rate of galaxies sSFR = SFR/M_* keeps rising with increasing redshift at $0 < z < 7$ ([426, 108, 436]). This indicates that, on average, galaxies are more vigously star forming at higher redshift. This monotonic rise in specific star

formation rate is consistent with the theoretical picture of galaxy formation fed by cold streams where the gas accretion occurs more vigorously at higher redshift (e.g. [110]).

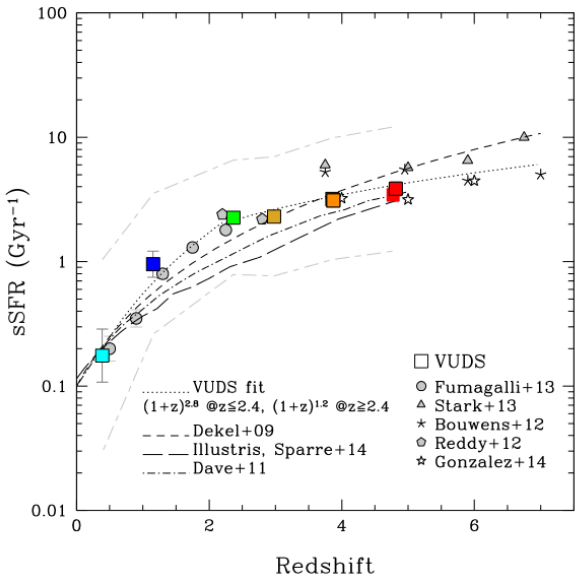

Figure 1.7: The evolution of the average specific star formation rate with redshift. The square points show the result from the population synthesis analysis of $i_{\rm AB}$-selected galaxies in VIMOS Ultra Deep Survey ([436]).

The population of galaxies is very diverse and the demographics evolve from low to high redshifts. It is known that the stellar content of galaxies exhibits the colour bimodality in the local universe $z \sim 0$, where the population can be separated into actively star-forming galaxies and quiescient galaxies containing passively assembled old stars (e.g. [36] for review). Based on the rest-frame UVJ colour-colour diagram, K_s-band selected galaxies in COSMOS/UltraVISTA ([315], Figure 1.8) and ZFORGE ([448]) show that nearly $\gtrsim 70\%$ of massive galaxies with $M_* \gtrsim 10^{11}$ M$_\odot$ are quiescent whereas lower mass galaxies with $M_* \lesssim 10^{11}$ M$_\odot$ are dominated by the star forming population in the nearby universe at $z < 1$. With increasing redshifts, the fraction of quiescent galaxies monotonically decreases to $\sim 50\%$ by $z \sim 2$. At $z > 2.5$, the star forming galaxies start to dominate the entire population of galaxies at all stellar masses ([315]). Although the colour bimodality persists at $0 < z < 2$, the seperation becomes less clear at higher redshift. Given the UVJ diagram classification, there is however a noticeable fraction of quiescent massive galaxies as much as $\sim 20\%$ at $3 < z < 4$ ([315, 431]). This evolution of the galaxy demographics suggests that major population of high-redshift $z > 2$ galaxies are star forming. While there might be a minor population of quiescent galaxies up to $z \sim 4$, the galaxies at $z > 4$ and during the EoR are likely to be predominantly star forming ones.

A concordance hypothesis to drive the colour bimordality is AGN feedback that quench star formation in massive galaxies. In contrast, stellar feedback in lower mass galaxies is thought to 'regulate' the star formation activity. Because of the quenching, massive

quiescent galaxies contain only old red stellar population and the stellar mass is assembled by passively merging galaxies.

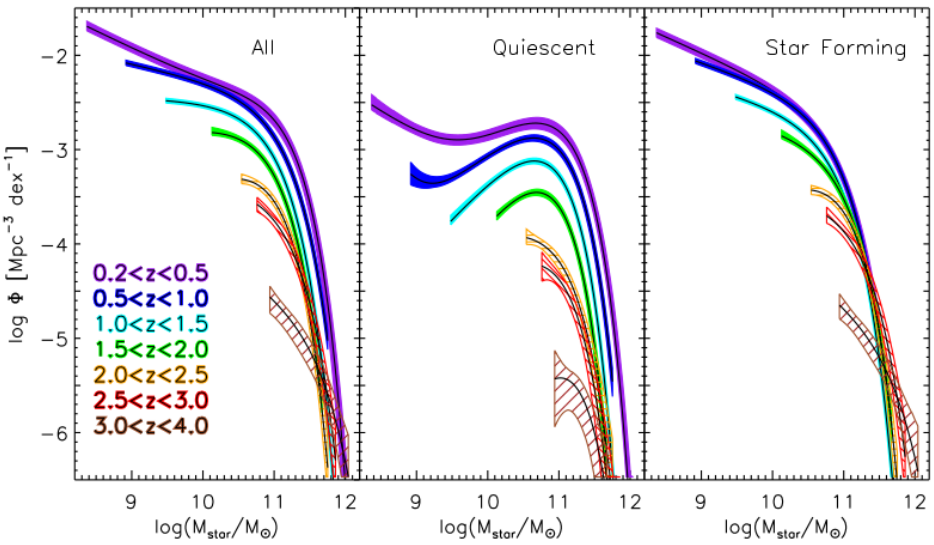

Figure 1.8: Stellar mass functions of all galaxies, quiescent galaxies, star-forming galaxies in different redshift derived from K_s-band selected samples in COSMOS/UltraVISTA ([315]).

Within a class of star-forming galaxies (SFR $\gtrsim 1-10$ M$_\odot$ yr^{-1}), two different modes of star formation are widely recognized, i.e. starburst and main sequence ([395] for review). This classification is based on the time variablity of star forming activity (star formation history) in galaxies. The former occurs in the bursty phase whose star formation timescale is much shorter than the age of galaxies (e.g. sSFR^{-1} $< 10^8$ yr at $z \sim 2$), whereas the latter forms stars gradually over the age of galaxies. At $z = 1.5-2.5$, the multi-wavelength survey data from COSMOS and GOODS fields suggests that the tight main sequence of star formation (Figure 1.9), which dominates the population of star-forming galaxies ([375]). Only 2% of the population presents starburst activity with a factor of $4-10$ boost in the star formation timescale ([375]). At higher redshifts $z \sim 3-6$, the average mode of star formation becomes more vigorous (Figure 1.7), which translates to the upward shift of the star formation main sequence in the vertical axis. With current data, it is still premature to make a statement on the relative importance between starburst and main sequence modes in star-forming galaxies at this epoch. However, based on the progressive analysis of the spectral energy distribution of galaxies ([387]), it is probably fair to say that $z \sim 3-6$ star-forming galaxies exhibit highly heterogenous physical properties in terms of star formation history and dust content etc.

I conclude this review by stating that present observations suggest that the major population of galaxies at $3 < z < 7$ is star forming at all stellar mass, along with a decreasing subdominant fraction of quiesent massive galaxies with increasing redshift at least up to $z \sim 4$. Therefore, as probably $> 80\%$ of the population at $z > 4$ is star foming galaxies, a large fraction of galaxies can be studied by the UV-selected samples

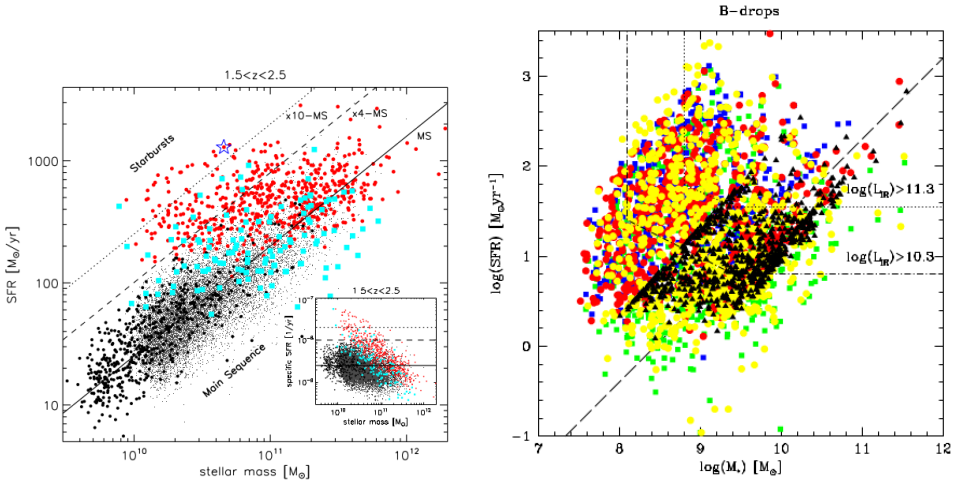

Figure 1.9: (Left) $1.5 < z < 2.5$ star-forming galaxies in COSMOS and GOODS fields [375] (Right) $z \sim 4$ LBGs. The different colours represent the difference in the star formation histories assumed in SED fitting analysis [387].

(which needs to be complemented by submm observations which provide dust-obscured star-forming galaxies missed by UV selection).

1.8.2 Lyman-break galaxies, Lyα emitters, Lyα emitting galaxies

Lyman break selections and narrow band Lyα selection are two primary methods to select high redshift galaxies, which have been successfully applied to study galaxies at $2 < z < 8$. We first set the terminology to avoid a confusion about what I mean by Lyman break galaxies (LBGs), Lyα emitters (LAEs), and Lyα emitting galaxies. LBGs are galaxies that are selected by Lyman break features of the spectra (a.k.a. dropout selection). LAEs are galaxies selected by narrow-band imaging targeting Lyα emission line. Lyα emitting galaxies refer to all galaxies that are emitting Lyα regardless of the selection methods. Both LBGs and LAEs thus refer to an observational classification of galaxies based on the selection technique, while Lyα emitting galaxies refer to the physical (spectral) properties of galaxies.

It is now viewed that these two class of samples, LBGs and LAEs, are essentially the same class of objects ([106]); both are high-redshift actively star forming galaxies although they could differ in terms of their stellar mass, age, and dust content.

1.8.3 Luminosity functions of LBGs and LAEs

The UV luminosity function of LBGs is steadily built up from $z \sim 8$ to $z \sim 4$ (see Figure 1.10 e.g. [53] and references therein), which is understandable from the hierarchical build-up of star forming galaxies. On the other hand, the Lyα luminosity function of LAEs between $z \simeq 3$ and $z \simeq 5.5$ is approximately constant, and shows a decline from $z \sim 5.7$

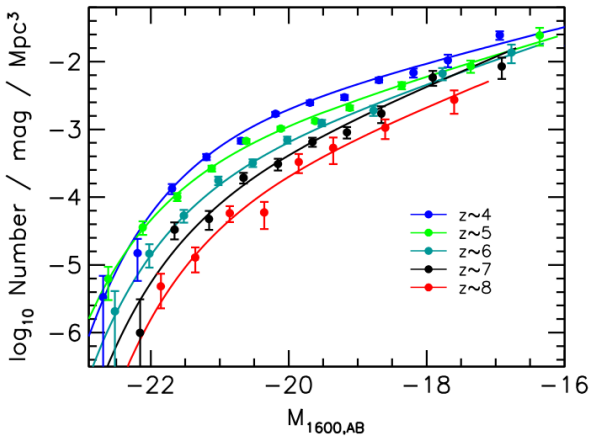

Figure 1.10: $4 < z < 8$ UV luminosity function of LBGs in HST Legacy Fields [53].

to $z \sim 7$ (Figure 1.11, [324, 65, 234]). The faint end slope,, which characterizes the abundance of lower mass galaxies, steepens with increasing redshift for both LBGs and LAEs. Physical origins of the steepening faint-end slope can vary from the preferential increase of the lower mass haloes hosting faint galaixes at higher redshift as a result of hierarchical structure formation to the evolution of star formation activity at a fixed mass due to changing gas accretion and feedback efficiency. The bright end of the UV luminosity of LBGs from $z \simeq 7$ to $z \simeq 5$ might also steepens with decreasing redshift, which could be interpreted as an indication of the onset of mass quenching (possibly AGN feedback) in high redshift galaxies ([55, 54]). A concordance view for the approximate non-evolution of Lyα luminosity function of LAEs at $3 < z < 6$ is because Lyα emission is controlled by the opposing evolution of dust content and the number density/luminosity of star forming galaxies. Hayes et al. ([189]) estimated that the average escape fraction of Lyα photons from galaxies is increasing as $f_{esc}^{Ly\alpha} \propto (1+z)^{2.5}$. This can compensate the decrease of the number density of galaxies at higher redshift expected from the simple hierarchical build-up. To back up this further, the UV luminosity function of LAEs is shown to grow from $z \simeq 5.7$ to $z \simeq 3$ unlike the Lyα luminosity function ([324]). This supports the view that Lyα emission from galaxies indeed has a complex dependence not only on the stellar assembly of star forming galaxies in the hierarchical structure formation, but also on the physics of dusty ISM, CGM/IGM, and radiative transfer in and around galaxies.

1.8.4 Lyα emission and absorption in LBGs: the physical state of high-z star forming galaxies

A subset of LBGs shows Lyα line profile in emission whereas other show Lyα line profile in absorption ([386] for a brief overview). This is known from the stacked spectra of LBGs at $z \sim 3$ ([400, 427]) as well as the SED analysis of photometric catalogues at $z \sim 3-6$ ([108]). More homogenous magnitude-limited samples from VIMOS Ultra Deep

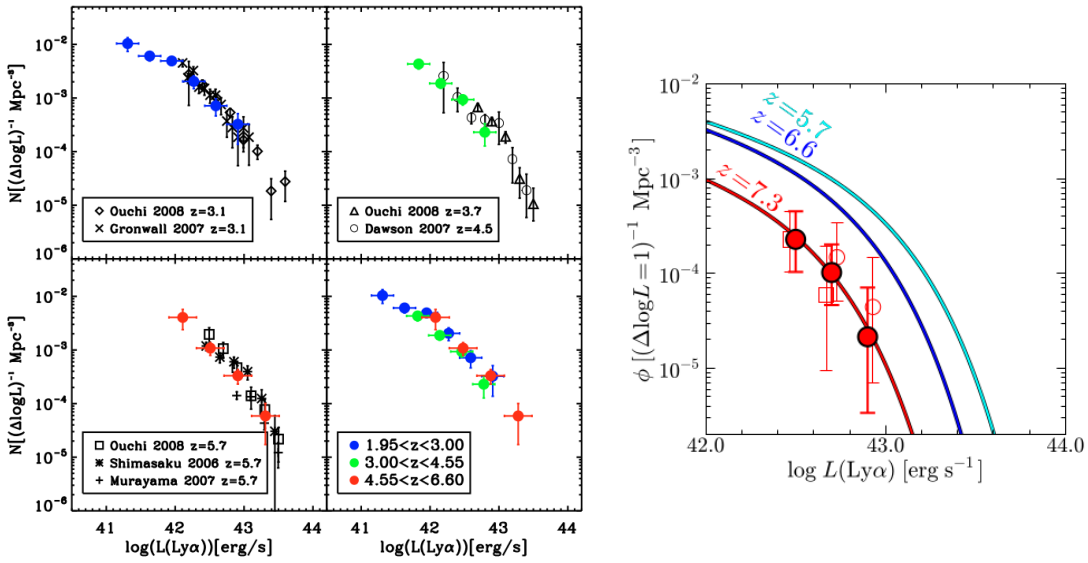

Figure 1.11: (Left) $3.1 \lesssim z \lesssim 5.7$ Lyα luminosity functions of LAEs in spectroscopic VIMOS VLT Deep Survey are shown by coloured symbols. The results from narrow-band surveys are shown by black symbols. (Figure taken from [65]) (Right) $5.7 < z < 7.3$ Lyα luminosity functions of LAEs from the narrow band survey with Subaru telescope ([234]).

Survey clearly confirm the variety of Lyα line profiles both in emission and absorption at $2 < z < 6$ ([246]). The follow-up spectroscopy of dropout samples at $3 < z < 7$ reveals that a subset of higher-z LBGs indeed shows Lyα in emission (e.g. [425]). Physically, the prevalence of Lyα line in LBGs is interpreted as a result of the different physical state of high-z star forming galaxies and their surrounding IGM.

The observed fraction of LBGs showing Lyα emission increases towards higher redshifts from $z = 2$ to $z = 6$, then declines beyond $z \gtrsim 6$ (see Figure 1.12, [117] and references therein). For $z = 2-6$, the decreasing dust content in LBGs on average towards higher redshift can elevate the detectable fraction of Lyα photons from galaxies. Bluer UV continuum slope, i.e. less dust reddening, on average at higher redshift supports this interpretation (e.g. [35]). Furthermore, when the parent LBG samples are divided into UV faint and UV bright objects, UV faint LBGs show a higher fraction of Lyα emission than UV bright LBGs at $3 < z < 6$ (e.g. [425, 321]). This is a consequence of the anti-correlation between Lyα equivalent width (EW) with UV magnitude, i.e. UV fainter objects have a higher probability to show larger Lyα equivalent width (e.g. [9, 463]). A physical interpretion is to invoke the increasing dust content for UV bright LBGs, which absorbs Lyα photons ([465]). This dust hypothesis for Lyα line formation works if the dust distribution within galaxies preferentially suppresses the escape of Lyα photons than the escape of continuum UV photons to boost Lyα EW. The UV continuum slope is bluer for objects with stronger Lyα emission among LBGs ([400, 425]), whereas LBGs with redder UV continuum slope show no or Lyα in absorption ([88]). As the UV slope correlates with UV luminosity ([131]

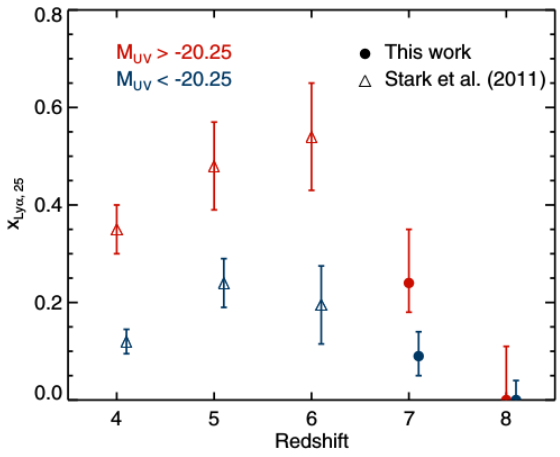

Figure 1.12: Fraction of LBGs showing Lyα emission with EW $\geq$ 25 Å. The parent LBG samples are divided into UV bright objects ($M_{\rm UV} < -20.25$) and UV faint objects ($M_{\rm UV} > -20.25$). [392]

and references therein), the dust hypthesis in controlling Lyα emission from galaxies is consistent with observations to explain the observed trends of the Lyα properties of LBGs as a function of both UV magnitude and redshift at $3 < z < 6$.

In addition to dust, other physical mechanisms such as the covering factor of neutral hydrogen and the kinematics of the ISM will be required to fully explain the observed Lyα properties of LBGs at $3 < z < 6$. In fact, it is known that dust is not only a factor controlling the fate of Lyα photon escape because a large variation of Lyα escape fraction from 0.1% to 100% exists in spectroscopic samples of galaxies even with low dust obscuration $E(B-V) < 0.05$ ([66]). Outflow and ionized channels in the ISM around stellar H II regions and galactic winds can help the escape of Lyα photons into the IGM. This means that a population of LBGs with outflow and a low covering factor of neutral hydrogen could show more prevalent Lyα line in emission. Observationally, the covering factor and velocity of neutral hydrogen in galaxies can be traced by the low-ionization ISM absorption lines ([428, 216]). Lyα equivalent width is larger for smaller equivalent width of the ISM absorption lines ([400, 216]) and the blueshifted velocity offset of absorption peaks, relative to the rest-frame systematics velocity of galaxies, indicates the presence of outflow in LBGs ([428, 479]). Although it is still not clear what physcial mechanism is really controlling the relation between the Lyα line formation and UV luminosity of LBGs, it is likely that the answer lies in the relation between star forming activity and the state of dusty turbulent ionized/neutral ISM and outflow of gas around star forming galaxies.

1.8.5 Clustering properties of LBGs and LAEs: host dark matter haloes and galaxy environment

Clustering analysis of LBGs and LAEs has been performed at a wide range of redshifts from $z \sim 2$ to $z > 6$. The measurements range from early results at $z \sim 2-3$ ([3, 31]) to recent results at $z \sim 3-7$ using early Subaru HSC data $z \sim 4-7$ ([186]), CFHTLS data and GOODS fields at $z \sim 3-5$ ([191, 248]), and the high-z data at $z > 6$ ([280, 326, 16]). These angular clustering measurement of LBGs has established that the host halo mass of LBGs is $\sim 10^{11.5} - 10^{12}$ $M_\odot$, whereas LAE clustering analysis shows that LAEs are preferentially located in smaller halo mass $10^{10} - 10^{11}$ $M_\odot$ ([235, 168, 325, 32]). Moreover, UV bright LBGs are more biased than UV faint ones, which indicates that the host halo mass is larger (e.g. [186]). Larger halo mass of LBGs than LAEs are not surprising given that SED fitting analysis also indicates that larger stellar mass for LBGs than LAEs on average.

Clustering analysis also gives an insight into the halo occupation number and duty cycle of LBGs and LAEs. The information on the halo occupation number comes from the small-scale clustering (one-halo terms). The excess of clustering signal at scales smaller than the radii of haloes indicates that a single parent halo contains multiple galaxies. For LBGs, such small-scale clustering is observed from $z \sim 3$ to redshift as high as $\sim 5-7$ ([219, 186]), whereas LAE studies have not convincingly shown the presence of multiple LAEs in a single halo (although the small-scale excess of clustering signal relative to the power-law fit seems to be present at $z \simeq 3$ ([325, 32])).

In addition, a comparison of the observed number density of LBGs and LAEs to the expected number of host haloes (whose mass is constrained from the large-scale clustering signal) constrains the duty cycle $f_{\rm duty}$, i.e. the fraction of haloes hosting LAEs and LBGs. LAEs at $z \sim 3-6$ seem to favour a duty cycle below 100% (as low as $\sim 1\%$) at $z \sim 3-6$, although with large uncertainties ([168, 235, 325]). Because not all star forming galaxies show observable Lyα flux in emission (§ 1.8.4), it is natural that only a subset of haloes at fixed host mass of star forming galaxies is selected as LAE hosts. LBGs also permit a duty cycle of tens of percent level, although $f_{\rm duty} = 1$ is consistent with data ([186]). Because $f_{\rm duty} < 1$ of LBGs may suggest another population of galaxies at high redshift[6], we must proceed carefully with improved clustering measurements before we draw a conclusion.

Cooke et al ([89]) suggested that the galaxy environments of Lyα emitting and Lyα absorbing LBGs at $z \sim 3$ are different; Lyα absorbing LBGs reside in group-like environments hosting multiple LBGs whereas Lyα emitting LBGs are located in the outskirts of galaxy groups and in the fields. They applied an interesting technique to separate the photometric samples of LBGs into Lyα emitting and absorbing objects using the fact that Lyα emitting/absorbing LBGs segregate in colour-magnitude diagram ([88]). The auto- and cross-clustering analysis of Lyα emitting/absorbing LBGs showed the excess clustering of both class of LBGs relative to the full parent LBGs as well as a sign of the anti-correlation

[6]For example, $f_{\rm duty} < 1$ of LBGs may be reconciled if there is a non-negligible fraction of quiescent galaxies or dusty starburst galaxies, which are missed by the Lyman-break selection.

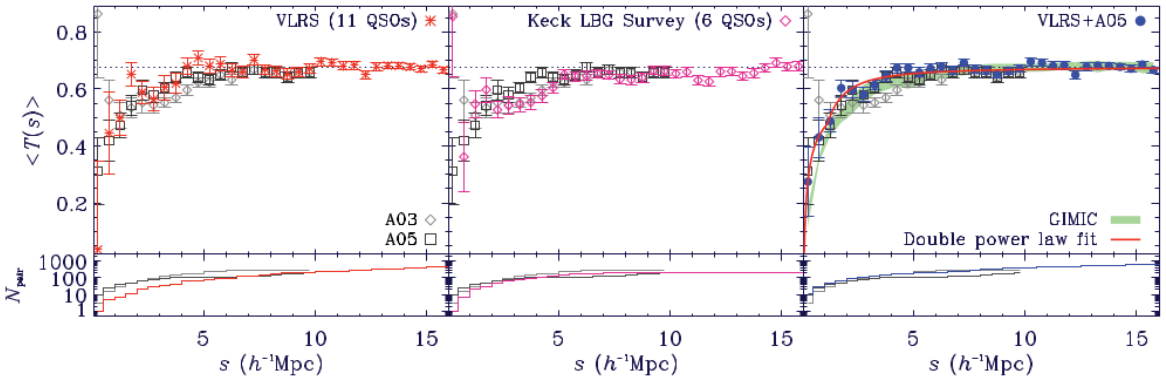

Figure 1.13: The mean Lyα transmitted flux around LBGs at $z \approx 3$ galaxies as a function of distance in the redshift space. The left panel shows the result from VLRS, the middle panel shows the result from Adelberger et al 2003 (A03) and 2005 (A05), and the right panel shows the combined VLRS+A05 result. The red solid curve shows a double power-law fit and the green shaded area is the result from a cosmological hydrodynamical simulation. (Figure taken from [455])

between the Lyα emitting and absorbing LBGs. Although a future confirmation with spectroscopic samples of Lyα emitting/absorbing LBGs is needed (which is potentially possible with e.g. VIMOS Deep/Ultra-Deep Survey ([246])), it suggests that the colour-density relation could be in place already at $z \sim 3$ similar to the local universe ([129], see also [358]). Extending this study to higher redshift is important to probe reionization. The possible relation of the galaxy envrionment with the clustering of Lyα emitting LBGs (as claimed by [89]) may be confused with the true reionization signal, which also causes the excess clustering of Lyα emitting LBGs relative to the full parent LBG samples.

1.8.6 Circum-/intergalactic gas environment of LBGs and LAEs

The intergalactic gas structure around galaxies can be directly studied by the observational strategy that surveys galaxies in the foreground of QSOs. In this strategy, together with the traditional spectroscopic/photometric study of galaxy properties, QSO absorption spectra are used to probe the physical state of intervening gas around galaxies. The seminal paper by Adelberger et al. in 2003 ([4]) has demonstrated the capability of such galaxy surveys in QSO fields. There are currently two large spectroscopic galaxy surveys in QSO fields at $z \sim 2-3$; Keck Baryonic Structure Survey (KBSS, led by C. C. Steidel) and VLT LBG Redshift Survey (VLRS, PIs: T. Shanks and L. Infante).

The clustering and kinematics of gas around LBGs are studied using three different, but closely related, statistics: (1) mean Lyα transmitted flux around galaxies ([4, 3, 92, 455]) (2) cross-correlation between galaxies and absorbers identified by Voigt profile fitting ([379, 480, 404, 378, 438]) (3) pixel optical depth method (median optical depth around galaxies) ([362, 456]). Although the three statistics differ in the method for charaterizing the gas in QSO absorption spectra, they provide a consistent physical picture about the

intergalactic gas environment around galaxies.

LBGs are on average associated with the overdensity of the intergalctic gas, but large rms fluctuations of Lyα absorptions suggest that the gas environment may consist of clumps and filaments. Figure 1.13 shows the mean Lyα trasmitted flux around LBGs at $z \sim 3$ measured by the VLRS ([92, 455]) and Keck-based surveys ([4, 3]), which clearly shows the enchanced mean Lyα absorption at $< 5h^{-1}$cMpc around the galaxies. Other statistics also agree with this picture. The number and column density of Lyα absorption lines ([378]) and the median Lyα optical depth ([362, 456]) also increase around LBGs. The large scatter around the mean Lyα transmitted flux ([3]) and median Lyα optical depth ([362]) suggest that despite the average increase of gas density near LBGs, a noticeable fraction of the sightlines around galaxies can show high Lyα trasmission. This supports the idea that the gas structure around galaxies is highly inhomogenous.

The redshift-space anisotropy of the clustering between LBGs and Lyα absorption has been observed in Lyα transmitted flux ([3, 455]), optical depth ([362, 456]), and absorber lines ([378])). Figure 1.14 shows the redshift-space anistropy of the median pixel optical depth around galaxies ([456]). Such redshift-space anisotropy can be physically caused by the gas flow around galaxies such as the large-scale infall (as in the Kaiser effect) and outflows. With a progressively improving determination of galaxy redshifts[7], the signatures of outflow and large-scale infall of gas onto galaxies have detected in the redshift-space distortion of the pixel optical depth-LBG clustering ([456, 457] see also [455]).

Metal absorption lines (high-ionization systems such as C IV and Si IV and low-ionization systems such as O I, C II and Si II) in QSO absorption line spectroscopy also contain viable information about the metal enriched gas environment around galaxies. At $z \sim 3$, C IV absorptions correlate with LBG overdensities ([4, 3, 92]), albeit with less than unity covering factor ([456]). The less tan unity covering factor supports the picture that the metal-enriched gas around the galaxies has an inhomogenous structure. Similarly to H I Lyα absorption, the signature of metal enriched gas flow around galaxies appears in the redshift-space anisotropy of LBG-C IV clustering as well as O II, C III and Si IV. The larger line-of-sight stretch of the finger-of-god effect in metal absorption lines than in Lyα is suggestive of galactic outflow or turbulent kinematics of metal enriched gas in galaxies.

Recently, spectroscopic galaxy survey in QSO fields has been pushed near the EoR for the first time[8]. Díaz et al. ([114, 115]) reported the distribution of LBGs and LAEs around $z \sim 5.7$ C IV absorption systems in two background QSO fields ($z_{\rm QSO} = 6.01$ and 6.31). C IV absorption at $z \sim 5.7$ appears to reside near LAEs rather than LBG overdensity, in contrast

[7]One must be careful about the spectroscopic redshift determination of LBGs, which mostly depends on Lyα emission lines and the interstellar absorption lines. Relative to the true galaxy systematic redshifts measured from the nebular emission lines (e.g. Hα, [O II] λ3727), the Lyα line is known to have a tendency to be redshifted due to radiative transfer effect and the interstellar absorption lines are blueshifted due to galactic outflow. These effects must be calibrated carefully to quantify the physical origin of the redshift-space anisotropy ([361]).

[8]There has been earlier imaging observations of high-redshift QSO fields to quantify the clustering of galaxies around the QSOs. For example, Stiavelli et al. conducted HST imaging observations of galaxies (i'-dropouts) in several $z \sim 6$ QSO fields ([430, 225]).

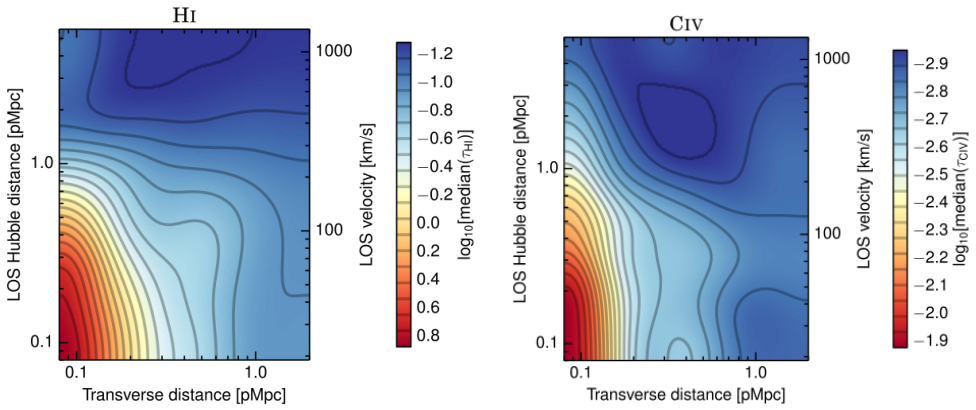

Figure 1.14: Two dimensional median pixel optical depth maps (H I and C IV) around LBGs at $z \approx 2.4$. The line-of-sight elongation is likely to be caused by the gas peculiar velocity around the galaxies. On ~ 1 pMpc scales, the compression along the line-of-sight in H I is indicative of large-scale gas infall onto the galaxies. (Figure taken from [456])

to the lower redshift $z \sim 3$ result where C IV absorption is strongly correlated with LBG overdensity. Accepting this result implies that high-redshift C IV systems near the EoR trace low-to-intermediate density environments dominated by low-mass galaxies ([114]). The interpretation of metal line diagnostics is complicated because it depends both on the chemical enrichment history by stellar nucleosynthesis/galactic winds and the ionization state of the CGM/IGM around galaxies ([150], see also [15, 20, 282]). Nonetheless, a tentative interpretation is that as LAEs (low-mass galaxies) dominate the total ionzing photon budget than LBGs (massive galaxies), the region near the overdensity of low-mass galaxies is more highly ionized. Such high ionization near LAEs brings high-ionization C IV systems to a detectable level ([114]). Despite the uncertanty in interpratation, the work of Díaz et al. demonstrated that spectroscopic galaxy survey in QSO fields near the EoR is a powerful approach to gain insight into the relation between the high-z galaxies and the physical state of the IGM during reionization.

1.8.7 State of the IGM at $2 < z < 5$

The physical state of the IGM in the post-reionized universe at $2 < z < 5$ places an important boundary condition that any successful model of reionization must satify.

The observed features of 1D Lyα forest flux power spectrum and flux probability distribution function measured by Keck/HIRES ([365, 97, 274]), VLT/UVES ([229, 226]), SDSS ([275]), and BOSS ([330]) can be captured remarkably well by the prediction from cosmological hydrodynamical simulations only with an additional model of the homogenous UV background. This suggests that the physical properties (density, ionization, and thermal state) of the low-density photoionized IGM at $2 < z < 4$ is largely controlled by the gravity+hydrodynamical process in ΛCDM cosmology and the metagalactic UV background from galaxies and AGNs, but insensitive to detailed model of feedback. There seems no

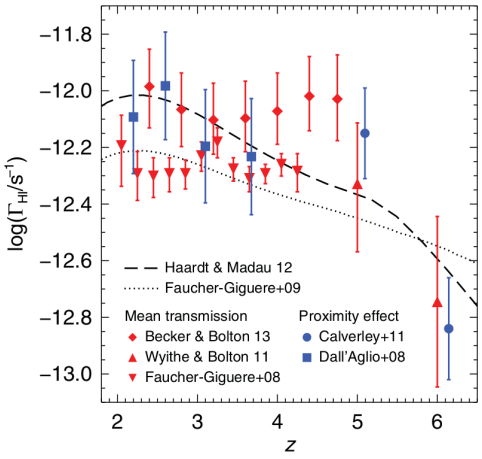

Figure 1.15: The recent measurements of H I-photoionization rate as a function of redshift. (Figure taken from [20])

missing physical process (as summerized in § 1.6.2) to understand the large portion of the IGM. Feedback plays only a secondary role (but noticieble with current statistical error of observations). Astrophysical effects such as self-shielding ([276]), galactic winds ([95, 112, 276, 468]) fluctuations of the ionizing background ([94, 294, 276]) and thermal structure from He II reionization ([244, 283]) seem to have small impact (about $\lesssim 10\%$) on the 1D flux power spectrum in 10^{-3} s km^{-1} $< k <$ 10^{-1} s km^{-1} range currently probed. 3D flux power spectrum measurable from dense samples of QSO spectra can be used to gauge the large-scale flutuations of the UV background and the thermal structure of the IGM ([283, 352, 178]).

A traditional Voigt profile fitting analysis views QSO absorption spectra as a superposition of individual Lyα absorption lines, which are characterizd by the column density $N_{\rm HI}$, the line width parameter b, and redshift (see [364, 295] for review with a historical development). According to the column density of a line, we can classify the absorption lines into Lyα forest absorbers tracing the low-density photoionized IGM and the self-sheilded gas such as Lyman-limit systems (LLSs) and damped Lyα systems (DLAs)[9] ([483] for review). The abundance of the absorbers is summerized in the H I column density distribution (CDDF). The shape of the CDDF evolves very slowly with redshift. To first order, the shape is well captured by a single power-law $\propto N_{\rm HI}^{\beta}$ where $\beta = -1.43 \pm 0.03$ in $12 < \log_{10} N_{\rm HI}/({\rm cm}^{-2}) < 21.5$ ([228]). Modern data indicates the bend in the CDDF ([320]). The agreement with hydrodynamical simulations strongly suggests that this bend is caused by the transition from the photoionized gas to self-shielded systems ([359]). Furthermore, the redshift evolution of the absorption line systems tells us about the neutral gas reservoir in the universe. The number density of DLAs stays approximately constant over $2 < z < 5$ with mild increase by a factor of two from $z = 2 - 4$ ([355, 356, 318, 93]),

[9]Although LLS and DLA are thought to come from a vicinity or as a part of galaxies, because these two objects are identified in QSO spectra, they are usually discussed in the context of IGM study.

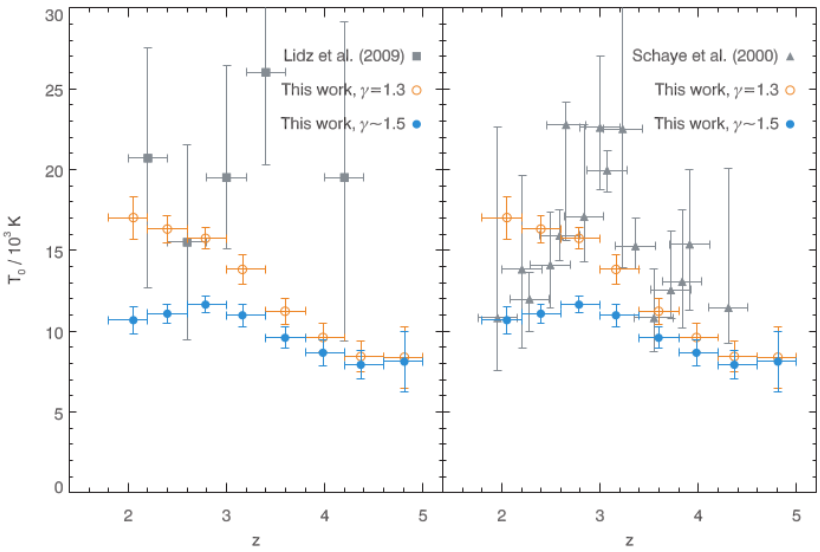

Figure 1.16: The temperature of the IGM at mean density as a function of redshift. Temperature measurements from curvature statistics (blue and yellow circles, [19]), wavelet analysis (gray squares, [252]), and Voigt profile analysis (triangles, [391]) are shown. (Figure taken from [19])

whereas the number density of LLSs ([347, 416]) and Lyα forest absorbers ([228]) show a faster increase with increasing redshift. The abudance and spatial distribution of absorbers set the neutral hydrogen cosmological mass density $\Omega_{\rm HI}(z)$ and the topology of neutral gas distribution at $2 < z < 5$ as well as controlling the scale of UV background fluctuations. This means that absorption line systems provide a valuable boundary condition for reionization scenarios.

The metagalactic UV background from star forming galaxies and AGNs shapes the ionization structure of the IGM after the reionization epoch. Many authors performed measurements of the H I-photoionization rate $\Gamma_{\rm HI}$ by matching the simulated Lyα forest, $\tau_\alpha \propto \Gamma_{\rm HI}^{-1}$, to the observed mean transmitted flux ([293, 46, 43, 141, 18]) and by the quasar proximity effect ([394, 101, 63]) (see Figure 1.15). These measurements show that the photoionization rate stays around $\Gamma_{\rm HI} \sim 10^{-12}$ s^{-1} at $2 < z < 4$ and declines rapidly at $z \simeq 5-6$. The photoionization rate depends both on the ionizing photon emissivity $\dot{n}_{\rm ion}(z)$ and the mean free path of ionizing photons $\lambda_{912}(z)$ as $\Gamma_{\rm HI} \propto \dot{n}_{\rm ion}(z)\lambda_{912}(z)$ (e.g. [295, 18]). In fact, the rapid decline of $\Gamma_{\rm HI}$ at $z \gtrsim 5$ is more likely caused by the shortening of the mean free path due to the increase in LLSs than by the rapid change in the ionizing photo emissivity. The total ionizing photon emissivity $\dot{n}_{\rm ion}(z)$ can be inferred from the measurements of $\Gamma_{\rm HI}$ and λ_{912}. The recent analysis by Becker & Botlon suggests that there are $2-14$ ionizing photons per hydrogen photons available per Gyr to ionize the IGM ([18]). This is a factor of two higher than the previous estimate, which suggested "photon-starved scenario of reionization" ([43]). Therefore, reinoization may not be so much photo-starved than previously expected; nonetheless, there is not an awful lot of excess photons relative

to hydrogen atoms that can be wasted.

The temperature and the equation of state of the IGM have been measured from the width of Lyα absorption lines such as via b-parameters ([391, 371, 59, 273, 377, 39]) and wavelet analysis ([443, 442, 252, 167]), curvature statistics ([19, 37, 38]), small-scale suppression of Lyα flux power spectrum ([495]), and flux probability distribution function ([23, 48, 467, 62, 247]). Although the simple tight power-law relation is expected to break down near He II and H I reionization epoch ([447, 286, 292, 86]), the effective slope γ, $T = T_0\Delta_b^{\gamma-1}$, has been a practical workhorse (see § 1.6.2 for theoretical motivation). Figure 1.16 shows the thermal history of the IGM. The temperature at a mean density reaches ~ 7000 K at $z \sim 5$, and gradually rises to $\sim 1 - 2 \times 10^4$ K from $z = 4$ to $z = 2 - 3$. Although a large scatter exists, the temperature rise is interpreted as an evidence of He II reionization at $z \sim 2 - 4$. The temperature measurement before He II reionization is also a valuable constraint on H I reionization at $z > 6$ because the relic signatures of the timing of reionization and the spectral shape of ionizing sources are retained in the thermal history at $z < 6$ ([442, 41, 80]).

Finally, He II reionization provides an intermediate redshift ($z \sim 2 - 4$) laboratory of the reionization process and a boundary condition for the contribution of harder ionizing sources such as AGNs during H I reionization. For recent progress, Cosmic Origins Spectrograph on the HST installed in 2009 is playing an important role. He II Lyα forest and its Gunn-Peterson trough suggest that He II is nearly completed by $z \simeq 2.7$ (see [282, 295, 364] and references therein). After the completion, the optical depth ratio between H I Lyα forest and He II Lyα forest gives the ratio between H I- and He II-photoionization rates, from which one can infer the relative contribution from star-forming galaxies and AGNs to the UV background at $z < 2.5$. Although it is difficult to draw a robust picture yet, such analysis seems to suggest AGNs contribute about $\gtrsim 50\%$ of the total H I-ionizing background around $z \sim 2$ ([45, 289]). Further constraints on the spectral shape of the UV background can be obtained from metal absorption lines (see [295] and references therein). This helps us to infer the contribution from star-forming galaxies (which generally lack the He II-ionizing photons due to absorption by stellar atmosphere) and AGNs (non-thermal emission and high-energy phenomena generally produces harder spectrum).

1.9 Reionization Probes

Although the broad-brush picture of reionization was in place even in the 90's, the nature of how reionization and cosmic dawn have occured is still largely unknown. When did reionization occur? What is the precise history and morphology of the reionization process? What are the sources driving reionization? How did the first galaxies and black holes form? Answering these questions is a key goal for extragalactic astrophysics in the 21st century. Before we set out our own voyage into the frontier of very distant universe, we can learn from the lessons from previous explorers. In the following, we review the progresses in last decades.

I organise the following discussion in terms of the questions concerning the timing,

history, morphology, and sources of reionization. The reionization probes discussed here are CMB, QSO/GRB absorption spectra, surveys of Lyα emitting galaxies, and 21cm cosmology ([137, 255, 354, 159, 312]). For other probes of reionization such as cosmic infrared backgrounds and X-ray backgrounds, I refer readers to [137, 255], which also contain references to the literature.

1.9.1 When did reionization occur?

Early important clues for answering this question come from the two breakthough observations of CMB and QSO spectra.

A first evidence of reionization comes from the observation of CMB. The measurement of the Thomson optical depth τ against electron scattering of CMB photons by ionized gas provides an integral constraint on the amount of ionized gas between us ($z = 0$) and the CMB ($z \simeq 1100$),

$$\tau = \sigma_T \int_0^z n_e(z) \left| \frac{dl_p}{dz} \right| dz \tag{1.25}$$

where σ_T is the Thomson cross section, $n_e(z)$ is the physical electron number density, and $|dl_p/dz| = c/[H(z)(1+z)]$. The latest Planck result constrained $\tau = 0.066 \pm 0.016$ from temperature power spectrum+CMB lensing ([351]), which translates to $z_{\rm re} = 8.8^{+1.7}_{-1.4}$ assuming an instantaneous reionization. The value is consistent with the improved series of WMAP results, $\tau = 0.089 \pm 0.014$ and $z_{\rm re} = 10.6$ ([193]).

The second evidence comes from observations of high-redshift QSO spectra. The UV photons with rest-frame wavelength $\lambda < 1216$ Å emitted from QSOs are redshifted along the line-of-sight by cosmological expansion. The photons are absorbed by the neutral hydrogen by Lyα scattering against the Lyα line $\lambda_\alpha = 1216$ Å. The optical depth against the Lyα scattering is

$$\tau_\alpha = \sigma_\alpha \int dz \left| \frac{dl_p}{dz} \right| n_{\rm HI}(z) \varphi_\nu(z) \approx \frac{c \sigma_\alpha n_{\rm HI}(z)}{\nu_\alpha H(z)}, \tag{1.26}$$

where $\sigma_\alpha = 0.011{\rm cm}^2$ Hz is the Lyα cross section and φ_ν is the line profile. The second approximate equality is called Gunn-Peterson optical depth, which is derived by James Gunn and Bruce Peterson in 1965 assuming an infinitely narrow line profile ([182]). The measurement of Lyα optical depth therefore tells us the amount of neutral gas at each redshift. Figure 1.17 shows the observed QSO spectra at $5 < z < 6.5$, the measured Gunn-Peterson optical depth, and the derived neutral hydrogen fraction of the IGM. The Gunn-Peterson optical depth is increasing rapidly at $z \sim 5 - 6$, indicating the increasing neutral hydrogen fraction at this epoch. As the Lyα line is very optical thick, a neutral fraction of only $x_{\rm HI} \sim 10^{-4}$ is sufficient to cause saturated absorption in the spectra. Thus, the measurement of the Gunn-Peterson optical depth only probes the tail end of reionization.

The above two evidences support the picture that reionization is mostly completed by $z \simeq 6$, and the era of reionization will probably be in the redshift range of $z \sim 6 - 15$.

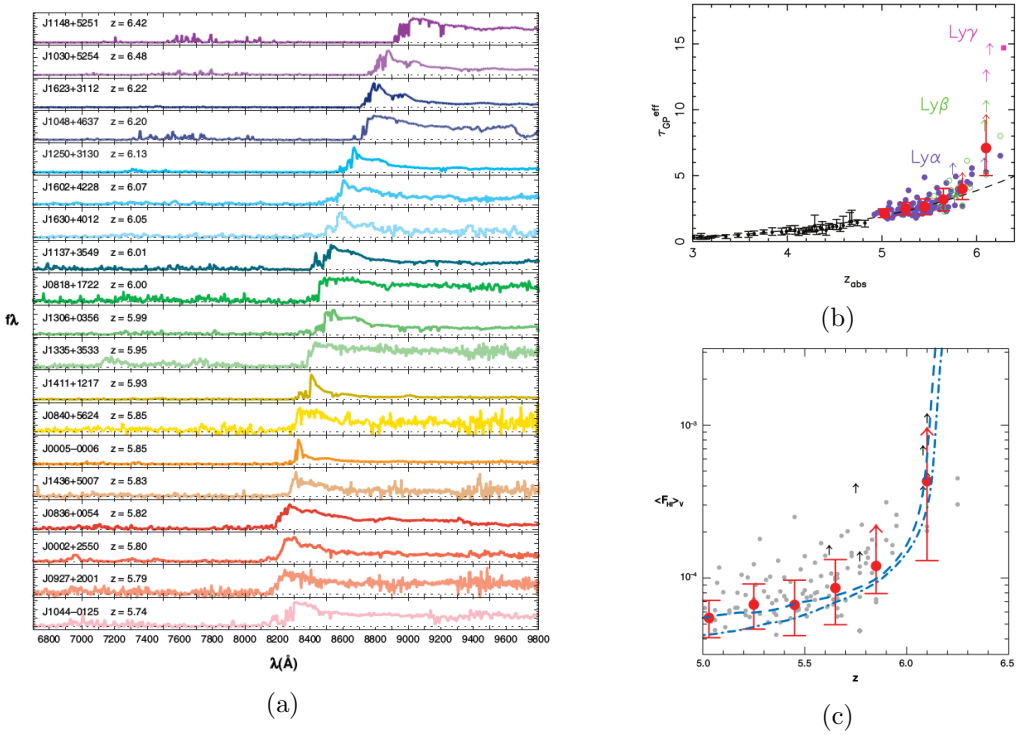

Figure 1.17: (a) High-redshift QSO spectra. (b) The Gunn-Peterson optical depth. (c) The neutral hydrogen fraction of the IGM. Figures taken from [139, 140].

However, exactly at what redshift the reionization has ended, or the ionization history, is not known yet. There is still room for incomplete reionization at $z \sim 5.5 - 6$ due to a possible presence of neutral patches of gas untouched by large H II regions ([254, 299]).

1.9.2 What is the history of reionization?

Most would agree that we yet do not have any robust constraint on the history and morphology of the reionization process. In other words, what is the global neutral (or ionized) fraction of hydrogen as a function of redshift, i.e. history? What is the size and distribution of H II regions during reionization, i.e. morphology? Recent progress starts to shed light on these two properties of the reionization process. However, these constraints are still very weak and subject of active research. We review constraints from (1) surveys of Lyα emitting galaxies, (2) QSO absorption spesctroscopy, and (3) CMB.

Surveys of Lyα emitting galaxies

Observing the high-redshift population of Lyα emitting galaxies is one way to probe reionization ([116, 117] for review). The idea is based on the fact that Lyα photons emitted from high-redshift galaxies are attenuated by the neutral gas along the lines-of-sight between the redshift of the population and us. Therefore, by having a way to infer the intrinsic Lyα

flux from galaxies unattenuated by the IGM, the difference between the observed Lyα flux and the intrinsic one tells us the amount of neutral hydrogen. The approach by interpreting the observed Lyα luminosity function of Lyα emitters using analytical models and/or simulations (e.g. [121, 214]) suggest the neutral fraction is of order of tens of per cent, $\langle x_{\rm HI}\rangle_V \sim 40\%$, at $z \sim 7$. The improved constraints by exploring the uncertainties from the assumed intrinsic galaxy properties and the detailed distribution of the neutral hydrogen in the IGM are still, interestingly, consistent with this value ([300, 77, 217]). If we accept this value inferred from Lyα emitting galaxies, by combined with Gunn-Peterson optical depth constraint, it suggests a rapid and late history of reionization where the neutral fraction decreases dramatically from $\mathcal{O}(10^{-1})$ at $z \sim 7$ to $\mathcal{O}(10^{-4})$ at $z \sim 6$.

QSO/GRB absorption spectroscopy

A more model-independent upper limit on the neutral fraction can be derived from studying the dark pixels, i.e. the region of QSO spectra with no transmitted flux ([299]). The fraction of dark pixels over all pixels at each redshift bin then translates directly to an upper limit on the neutral fraction. Because this involves no assumption on the reionization model, the upper limit can be considered as a fairly robust constraint on the reionization history. The updated upper limits derived from 22 QSO spectra are $\langle x_{\rm HI}\rangle_V \leq 0.06 + 0.05(1\sigma)$ at $z = 5.9$ and $\langle x_{\rm HI}\rangle_V \leq 0.04 + 0.05(1\sigma)$ at $z = 5.6$ ([278, 277]).

QSO near-zone spectra, the proximity of a QSO, show the transmitted flux. The interpretation of the Lyα damping wing feature in QSO near-zone provide constraints on the ionization state of the IGM. Analysing the near-zone spectra with the help of models or numerical simulations ([304, 393, 47]) then provides a constraint on the neutral fraction of the IGM. A comparison of the observed near-zone spectra with the mock spectra derived from simulations of patchy reionization placed a constraint that $\langle x_{\rm HI}\rangle_V > 0.1$ at $z \approx 6$ ([393]). A discovery of $z \approx 7.1$ QSO pushed the neutral fraction constraint to $z \sim 7$, placing the lower limit of $\langle x_{\rm HI}\rangle_V > 0.1$ ([47])[10]. We should keep in mind the uncertainties in these near-zone constraints, which come from the large line-of-sight variations ([221]) and a confusion with the presence of DLA along a QSO line of sight. Once we accept the QSO near-zone constraint, the stringent lower limit at $z \approx 6$ supports the picture of the late reionization history where substaintial diffuse neutral gas still remains at $z \approx 6.2$.

GRB afterglows provide constraints similar to those obtained from QSO near-zone. The high-redshift GRBs used as a probe of reionization are GRB 130606A at $z = 5.91$ ([76, 449, 450, 187]), GRB 050904 at $z = 6.295$ ([220, 451, 285]), GRB 140515A at $z = 6.33$ ([75, 296]), and GRB 080913 at $z = 6.7$ ([179, 335])[11]. An interpretion of the Lyα damping wing of GRB 050904 using radiative transfer simulations of reionization indicates that the

[10]However, recently the presence of damping wing feature was re-visited. It is argued that the feature may disappear when an improved estimate of QSO continuum is used ([50]). Therefore, the interpretation of the near-zone of $z \sim 7$ QSO must be taken with caution.

[11]The spectrum of GRB 090423 at $z = 8.26$ ([435, 381]) is too noisy to derive the neutral fraction from the damping wing. GRB 090429B ([98]) has photometric redshift of $z \sim 9.4$, but no spectroscopic observation is reported.

neutral fraction of $\langle x_{\rm HI}\rangle_V \approx 0.5$ at $z \simeq 6.3$ is consistent with the observed damping wing feature ([285]), although the constraint is weak due to the degeneracy with a possible DLA of the GRB-host galaxy. Nonetheless, unlike QSO spectra probing highly biased region of the IGM , GRB can probe an average region of the universe as it originates from the death of massive stars in galaxies.

CMB kinetic Sunyaev-Zel'dovich effect

Zahn et al. ([494]) provided an alternative route to constrain the reionization history using the kinetic Sunyaev-Zel'dovich (kSZ) effect of CMB ([369] fore recent review). Such constraint is useful as it is independent of Lyα transfer physics which underlies in both the interpretation of Lyα emitting galaxies and QSO/GRB absorption spectra described above. The Sunyaev-Zel'dovich effect causes secondary anisotropy and spectral distortion of the CMB imprinted by the scattering of CMB photons with free electrons after recombination. In particular, the kSZ effect is caused by the relative bulk velocity of ionized gas with respect to the CMB rest-frame (hotspot/coldspot when the ionized gas is moving towards/away from us). The data from the South Pole Telescope provided the kSZ constraint on the duration of the reionization process as $\Delta z \leq 7.9$ between $\langle x_{\rm HI}\rangle_V = 0.01$ and 0.8 ([494]).

Global 21cm signal

21cm line is produced by the hyperfine structure of the 1s state of a hydrogen atom. The abundant neutral hydrogen at $z > 6$ produces the redshifted 21cm signal at cosmological distance, which can then be observed at low radio frequency $1420/(1+z)$ MHz on the Earth. Because this signal directly traces the netural hydrogen, both the all sky 21cm signal (global 21cm signal) and the three dimensional imaging of the fluctuations (21cm tomography) are the most promising probes of reionization (see [262, 354, 159, 312] for details about 21cm cosmology).

Although 21cm signal has not been detected yet, the global 21cm experiment called EDGES using a single dipole antenna placed a (weak) lower limit on the duration of reionization, $\Delta z > 0.06$ ([57, 56]). This constraint is based on the undetection of a kink in the all-sky spectrum between 100 and 200 MHz ($6 < z < 13$). Such kink is expected if reionization occurs instantanously, causing a discrete jump in 21cm brightness temperature. While the constraint is not strong enough to narrow down the reionization history, it demostrated the potential of 21cm cosmology on observational basis.

The above four techniques using (1) Lyα emitting galaxies, (2) QSO/GRB absorption spectroscopy, (3) CMB kSZ effect, and (4) global 21cm signal summerize the current status of the endevour to map out the reionization history beyond the classical method using the CMB optical depth and Gunn-Peterson trough. Despite the impressive progress in constraining the neutral fraction of the IGM, we should keep in mind that there are still uncertainies both in observations and theoretical models as well as tremendous degeneracies

involved in the interpretation of the data. It is a goal of the coming decades to improve constraints on the reionization history to draw a robust conclusion on reionization.

1.9.3 What is the morphology of reionization?

The morphology of reionization, which is characterised by the size, shape, and distribution of cosmological H II regions, is an important property of the reionization process of an order higher than the one of the global ionization hisotry. We have a good physical motivation to think that reionization has a patchy morphology as predicted by cosmological radiative tranfer simulations since the turn of century (e.g. [172, 366, 82, 208] for earlier works). However, it is still difficult to draw a conclusion on the reionization morphology based solely on observational evidences. Various observations are consistent with and point to a patchy reionization picture, although they may not exclude other possibilities (e.g. homogenous reionization) with $> 5\sigma$ significance. I review observational clues about the morphology of reionization.

Surveys of Lyα emitting galaxies

The morphology of H II regions imprint its signature on the population of Lyα emitting galaxies. Lyα fraction test, i.e. a fraction of Lyα emitting objects in LBGs, potentially provides a way to study the distribution of H II regions around galaxies ([321, 343]). Observationally, a larger decline of Lyα fraction in UV faint galaxies than in UV bright objects for increasing redshift is interpreted as an evidence of patchy reionization ([321]). This is based on the picture of patchy reionization that UV bright galaxies is more likely located inside larger H II regions than UV faint objects. Theoretical studies support this interpretation, although the interpretaion is not unique to patchy reionization scenario ([217]). Another diagnostic is to study the distribution of Lyα emitting galaxies as a function of LBG overdensity, although the interpretion is still in infancy ([67]).

Clustering of Lyα emitting galaxies is a more direct signature of patchy reionization. During reionization, Lyα emission from galaxies is preferentially seen in those located in large H II regions, and contribute to the excess clustering signal over the gravitational origin from halo clustering ([284]). This excess clustering of reionization origin can be regarded as an evidence of patchy reionization. Subaru narrow-band surveys ([218, 325]) reported the measurement of angular clustering of LAEs at $z \simeq 6.5$. Current observations have not seen an obvious excess in the clustering amplitude at $z > 6$, although there are still large observational uncertainties. Also on the theoretical front, more quantitative studies of the signature of reionization, together with other effects such as duty cycle and host halo mass, are under way to full exploit the upcoming clustering measurements ([214, 213, 205, 413]).

QSO absorption spectroscopy

The scatter in the distribution of effective Lyα optical depth measured from QSO spectra at $4 < z < 6$ has provided an observational hint on the morphology of the reionization process.

The scatter in the effective Lyα optical depth is caused by the fluctuations of density, UV background, and temperature of the IGM. Becker et al. ([22]) found that while the observed scatter of the effective optical depth can be explained by the density fluctuations alone at $z \lesssim 5$, there is a larger scatter at $z > 5$. They interpreted the cause of this as a large UV background fluctuation at $z > 5$, which is a relic of the patchy reionization process. While there is an alternative interpretion as a signature of temperature fluctuation ([102]), both interpretations suggest a patchy reionization scenario.

In addition, the size distribution of dark gaps, the contiguous saturated regions in the Lyα forest in QSO/GRB spectra, provides an addtional diagnostic ([415, 334, 299, 164, 163, 165, 267]). Observations ([139]) show that the size of dark gaps increases with increasing redshift and rises rapidly near $z \sim 6$. At the same time, observations indicate the presence of large dark gap size $\gtrsim$ 30Mpc even at $z \sim 5.5$. The shape of the dark gap size distribution is expected to be different between the IGM during the patchy reionization process and the post-reionized IGM ([299, 267]). Therefore, the interpretion of this dark gap sizes could provides further insights on the end of reionization. The physical origin of the dark gaps, whether they originate from diffuse neutral gas between the H II regions or residual neutral gas in the highly photoionized medium with the UV background fluctuation, can be studied through the possible (non-)presence of damping wing in the stacked spectra around the dark gaps ([267]).

21cm power spectrum and tomography

Once the 21cm signal is convincingly detected observationally, 21cm cosmology is the most direct probe of the morphology of reionization. Recently the 21cm community has started to place an upper limit on the 21cm power spectrum (GMRT; [329, 328], PAPER; [6, 333], MWA; [125, 126]). LOFAR ([462]) has also reported deep radio observations in two EoR windows (North Celestial Pole field; [490], ELAIS-N1 field; [211]). Note that all the current 21cm experiments suffer from issues of calibration errors due to ionosphere, polarization leakage, and residual sidelobes of the point spread function, which contribute to systematic noise. Achieving accurate calibration is as important as reducing the thermal noise limit of the intruments. Arguably, progress in 21cm science in next few years will be driven by improving calibration techniques and foreground mitigation strategies. In the SKA-era, 21cm tomography will provide images of the IGM to probe the morphology of the reionization process directly.

1.9.4 What is the source of reionization?

Once the reionization history is constrained, a natural question is to ask about the sources of ionizing radiation. Do galaxies drive the reionization process? If so, what type of galaxies are responsible for reionizing the universe? Could QSOs contribute to the reionization process? What is the physical mechanism regulating the escape of ionizing radiation from galaxies to drive reionization?

The role of galaxies in driving reionization

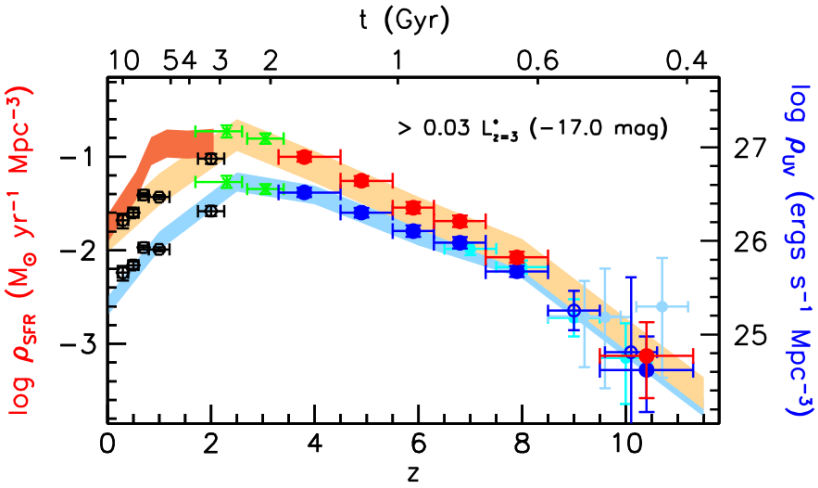

Figure 1.18: Cosmic star formation rate density and UV luminosity density as a function of redshift and cosmic time (Figure taken from [53]).

High-redshift surveys of galaxies are directly mapping out the population at $z > 5$. Figure 1.18 shows the cosmic star formation rate density $\rho_{\rm SFR}$. $\rho_{\rm SFR}$ is derived by integrating the observed UV ($\simeq$ 1500 Å) luminosity function down to a limiting magnitude ($M_{\rm UV} = -17$ in Figure 1.18), and coverting the UV luminosity to star formation rate based on the stellar population synthesis modelling. The dust attenuation must also be corrected (red symbols). Currently, the correction is based on the dust reddening effect of the UV continuum slope. However there is still large uncertainty associated with dust content of high-redshift galaxies. Despite these uncertainties, there is a growing consensus that the star formation rate density steadily decreases from $z \sim 4$ to $z \sim 8$ ([131, 148, 52] for review). Interestingly, there is an apparent steepening in the decline of the star formation rate density beyond $z > 8$. This may indicate the rapid emergence of star-forming galaxies in this epoch, although the observational uncertainties prevent us to draw any conclusion on the physical origin of this.

To address the role of the observed galaxy population in driving the reionization process, one can combine the observed SFR density and a simple analytical model of reionization,

$$\frac{dQ_{\rm HII}}{dt} = \frac{\dot{n}_{ion}(z)}{\bar{n}_H(0)} - \frac{Q_{\rm HII}}{\bar{t}_{rec}(z)}, \tag{1.27}$$

where $Q_{\rm HII}$ is the volume filling factor of H II regions. In a simple model of reionization, $Q_{\rm HII}$ is directly converted into the neutral fraction of the IGM, $\langle x_{\rm HI}\rangle_V = 1 - Q_{\rm HII}$. The ionizing emissivity $\dot{n}_{ion}(z)$ is inferred from the observed SFR density,

$$\dot{n}_{ion}(z) = f_{esc}\xi_{ion}\rho_{\rm SFR}, \tag{1.28}$$

where f_{esc} is the escape fraction of ionizing photons and ξ_{ion} is the rate of producing ionizing photons per unit star formation rate. The predicted reionization history is then compared to match the CMB Thomson optical depth and other constraints. This argument leads to a picture that if galaxies are a main reionizing agent, (i) the escape fraction must increase to the value as large as $\sim 20\%$ and (ii) there must be a sufficient population of faint galaxies below current detection limits. Despite the largely accepted view that galaxies are most likely contributors to reionization, we have not yet directly seen the population to draw any solid conclusion about the sources of reionization.

1.10 Conclusion

In this chapter, I have reviewed the foundations of observational cosmology, astrophysics of galaxies and the IGM, observations of high-redshift universe, and current status of reionization research. Given such foundations, we can now set out our voyage to test the current paradigm and to seek for the true nature of the EoR!

Chapter 2

Lyα Emitting Galaxies as a Probe of Reionization

In Chapter 1 we have learned the foundations of extragalactic astrophysics, which set the solid starting point to understand the Epoch of Reionization. In Chapter 2, I present the methods to constrain reionization using Lyα emitting galaxy surveys.

The chapter is organized as follows. Section 2.1 discusses backgrounds. 2.2 Section 2.2 briefly reviews the cosmic history of the H I content in the universe, ranging from the epoch of reionization to the post-reionized universe. In Section 2.3 we present our analytic framework of cosmological Lyα radiative transfer. In Section 2.4 we describe the methodology employed to generate the reionization models (bubble, web, and web-bubble models), as well as the intrinsic and apparent mock catalogue of Lyα-emitting galaxies. Section 2.5 shows our results. The conclusions and discussion about implications on Lyα-emitting galaxy surveys are then presented in Section 3.7.

2.1 Backgrounds

The Epoch of Reionization (EoR) and Cosmic Dawn are the least explored frontiers in observational cosmology and extragalactic astrophysics ([255]). Galaxy surveys are one of the most important pillars of modern cosmology, allowing us to study high-redshift galaxy formation and the reionization process of the intergalactic medium (IGM). Surveys of high-redshift galaxies using Lyman-break drop-out technique (Lyman Break Galaxies, LBGs) (e.g. [281, 138, 53]) and narrow-band filter targeting Lyα emission (Lyman Alpha Emitters, LAEs) (e.g. [265, 199, 325]) have provided a deep sample of objects, indicating that reionization requires many faint galaxies below the sensitivity limit of the surveys ([374]). Furthermore, observations of QSO spectra (e.g. [140, 22]) and the Cosmic Microwave Background (CMB) ([193, 351, 494]) offer hints that reionization is mostly completed at $z \gtrsim 6$.

However, beyond such indications, our present observational constraints on the EoR are still scarce, regarding both the reionization history and its topology/morphology. While

21cm experiments with radio interferometers such as LOFAR[1], MWA[2], GMRT[3], PAPER[4], HERA[5], SKA[6] offer the most direct probe of the physical state of the IGM during the EoR (e.g. [354]), the challenge in foreground removal and calibration remains. Substantial progress has been recently made by the 21cm community (e.g. [328, 333, 490]), but a detection is still missing. In principle, surveys of Lyα-emitting galaxies (Lyα surveys) offer an alternative and independent means from 21cm experiments to probe the EoR and constrain the global H I fraction (e.g. [116]). Such an approach is attractive because of the present availability of data and up-coming surveys with the Hyper Suprime-Cam on Subaru[7] and with future telescopes such as JWST[8], E-ELT[9], TMT[10], and GMT[11]. Furthermore, using multiple independent strategies can provide constraints on reionization which are less sensitive to systematic uncertainties of individual probes.

The challenge in using Lyα emitting galaxies as a probe of reionization lies in correctly interpreting observations ([321, 99, 425, 452, 64, 445, 66]). The reduced visibility of Lyα emission from galaxies at $z > 6$ has already been used to infer the global H I fraction of the IGM (e.g. [121, 213]). However, a robust interpretation is still uncertain because of the complex radiative transfer of both ionizing and Lyα photons. The Lyα transfer involves a wide range of scales including (*i*) the interstellar medium (ISM), where dust and gas distribution and kinematics determine the escape fraction of Lyα photons as well as their spectral line profile (e.g [466, 180, 206]); (*ii*) the circum-galactic medium (CGM), i.e. the direct environment of galaxies out to a few hundred kpc (e.g [120, 501, 502, 245]); (*iii*) the IGM, which - during reionization - contains diffuse neutral gas surrounding large ionized bubbles which themselves contain dense, self-shielding gas clouds. In order to obtain robust constraints on the global H I fraction, it is essential to understand the cosmological Lyα RT on all these scales.

The precise ionization structure of the IGM, i.e. the topology of reionization, is not only characterized by the size, abundance and distribution of large-scale ionized bubbles, but also by the small-scale dense H I absorbers self-shielded against the external ionizing sources. Interpretations of Lyα-emitting galaxies contain (often implicit) assumptions about the ionization structure of the IGM, mostly because of the difficulty to cover the entire dynamic range that is required to properly describe both the small-scale dense H I absorbers *and* the large-scale diffuse neutral IGM in reionization simulations. Two extreme assumptions, described in the first two bullets below, have been commonly adopted in the literature. Here we introduce the following terminology:

[1] http://www.lofar.org
[2] http://www.mwatelescope.org
[3] http://gmrt.ncra.tifr.res.in
[4] http://astro.berkeley.edu/~dbacker/eor
[5] http://reionization.org
[6] http://www.skatelescope.org
[7] http://www.naoj.org/Projects/HSC/
[8] http://www.jwst.nasa.gov/
[9] http://www.eso.org/public/teles-instr/e-elt/
[10] http://www.tmt.org/
[11] http://www.gmto.org/

- *Bubble model*: in this model small-scale H I absorbers are neglected. Under this assumption, the global H I fraction measures the H I content of the diffuse neutral IGM outside ionized bubbles.We refer to this as the 'bubble model'.

- *Web model*: here only the small-scale H I absorbers are considered. As this overdense gas largely traces the large-scale cosmic web, we refer to it as the 'web model'.

- *Web-bubble model*: reality is a combination of the two extreme configurations above. We refer to cases that contain both neutral phases (diffuse and clumped) of gas as the hybrid 'web-bubble model'. One can visualize this as the more common bubble model, but with 'impurities' in the ionized bubbles in the form of small-scale neutral islands.

Most previous works interpreting the observed reduction in Lyα flux from $z > 6$ galaxies have favoured a very high value of H I fraction, as high as $\sim 50\%$ at $z \sim 7$ (e.g. [284, 121, 214]). These studies used large-scale reionization simulations which did not have the spatial resolution to resolve the self-shielded small-scale absorbers.

The lack of self-shielding gas inside ionized bubbles in large-scale reionization simulations is clearly problematic: Lyα forest observations indicate that in the post-reionized universe, i.e. $z < 5$, H I gas is locked up in damped Lyα systems (DLA) and Lyman-limit systems (LLS) (e.g. [483]). Self-shielded absorbers (LLSs and DLAs) are also expected to reside inside ionized bubbles during reionization (and possibly with larger number densities, see e.g. [44]). The first investigations of hybrid web-bubble models have recently been reported ([300, 77]). Interestingly, these papers still favour large values for the H I fraction, as high as $\sim 40\%$ at redshift $z = 7$.

In this chapter we investigate the impact of large-scale patchy reionization and small-scale H I absorbers on the observed Lyα flux of galaxies, and its implication on the H I fraction measurements from Lyα surveys. We explore a unique combination of cosmological hydrodynamical, radiative transfer simulations and analytic models. Our analytic framework is powerful as it facilitates the interpretation of the results of our simulations, and provides us with a tool to quickly explore a large range of parameters describing the reionization and Lyα transfer processes in future work.

2.2 Cosmological H I content

In this section, we review the redshift evolution of the H I content both during and after reionization. This can be quantified either by the mass-weighted $\langle f_{\rm HI} \rangle_M$ or the volume-weighted $\langle f_{\rm HI} \rangle_V$ neutral fraction. A compilation of current estimates in the literature is shown in Fig. 2.1.

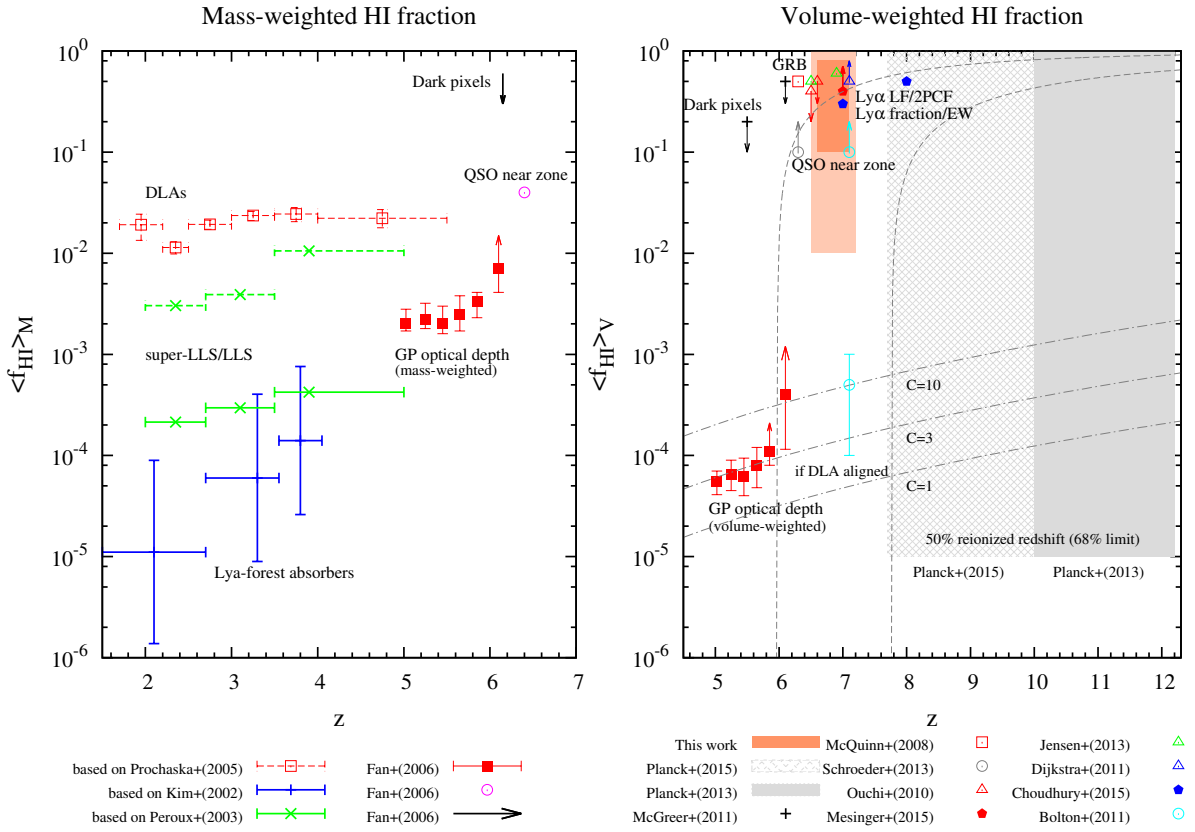

Figure 2.1: Cosmological H I fraction $\langle f_{\rm HI} \rangle_{V,M}$ in the diffuse IGM and high-column density Lyα absorbers (LLS/DLAs) at various redshifts, from the post-reionized universe to the epoch of reionization. The mass-weighted H I fraction embedded in the small-scale absorbers is computed from equation (A.2) using the fitting function to the observations of Prochaska et al. (2005), Kim et al. (2002) and Peroux et al. (2003). All the other values are compiled from the literature, as indicated by the labels above (z=7 constraints by Dijkstra et al. (2011) and Jensen et al. (2013) are shifted by +0.1 and -0.1 in redshift to avoid a cluttering of data points). The dash-dotted lines in the right panel are the volume-weighted neutral fraction of a diffuse IGM with clumping factor $C = 1, 3, 10$ in photoionization equilibrium with a UV background $\Gamma = 10^{-12}$ s^{-1}. The dashed lines are the bubble model estimates of the neutral fraction in H I patches outside ionized bubbles (left and right lines are DEC and CONST models in § 2.2.3).

2.2.1 Observational Constraints on H I in the Post-Reionization Epoch

The left panel of Fig. 2.1 clearly indicates that the post-reionized universe still contains neutral islands of gas in the form of self-shielding LLSs and DLAs. The abundance of the H I gas is generally quantified in terms of the H I column density distribution function (CDDF), $f(N_{\rm HI}, z)$, which is defined as (e.g. [483]) $f(N_{\rm HI}, z) = \frac{\partial^2 \mathcal{N}}{\partial N_{\rm HI} \partial z} \frac{H(z)}{H_0 (1+z)^2}$, where $\frac{\partial^2 \mathcal{N}}{\partial N_{\rm HI} \partial z}$ is the number of Lyα absorbers $\mathcal{N}(N_{\rm HI}, z)$ per unit H I column density $N_{\rm HI}$ and per unit redshift, $H(z) = H_0[\Omega_m (1+z)^3 + \Omega_\Lambda]^{1/2}$ and H_0 is the Hubble parameter today. The mass-weighted H I fraction embedded in small-scale absorbers is estimated from observations of $f(N_{\rm HI}, z)$ (see Appendix A.1)[12].

The left panel of Fig. 2.1 further shows the mass-weighted H I fraction embedded in each type of Lyα absorber. The dominant reservoir of H I gas is the high-column density Lyα absorbers, mainly DLAs. The $\langle f_{\rm HI} \rangle_M \sim 1\%$ embedded in DLAs stays approximately constant over $2 < z < 5$, while the H I fraction embedded in super-LLS and LLS, which is the second dominant H I gas reservoir, increases with redshift. The diffuse IGM, represented by the Lyα forest absorbers, is highly ionized and remains a minor reservoir of neutral gas.

2.2.2 Observational Constraints on H I During Reionization

In the right panel of Fig. 2.1 we have compiled various inferred values of the volume-weighed H I fraction available in the literature from CMB ([349, 351]), Gunn-Peterson optical depth ([140]), dark pixels ([278]), Gamma Ray Burst afterglow ([285, 451]), quasars (QSOs) near zone ([47, 393]), Lyα luminosity function, equivalent width distribution, Lyα fraction, and correlation function ([325, 121, 214, 300, 77]). We also show our suggested constraint using the Lyα luminosity function alone (faint orange box, see § 2.5.1) and when combined with the equivalent width distribution (darker orange box, see § 2.5.6).

All the open points are simulation(model)-calibrated measurements, which use the Lyα radiative transfer modelling in the IGM around galaxies and QSOs. While previous works make very sensible assumptions to interpret the observed data, astrophysical systematics in such simulation(model)-calibrated measurements may raise questions about the robustness of the inferred values. While the present estimates at $5 < z < 7$ favour a volume-weighted H I fraction as high as $\sim 50\%$ if taken at face value, it should be kept in mind that these estimates are implicitly assuming a bubble model.

Interestingly, recent constraints ([300, 77]) including both large-scale patchy reionization and small-scale absorbers still favour values for the H I fraction $\gtrsim 40\%$ at $z > 7$. Our work also prefers numbers in this range. It is the aim of this chapter to understand the reason for this.

[12]Converting $f(N_{\rm HI}, z)$ into a constraint on the volume-weighted H I fraction, $\langle f_{\rm HI} \rangle_V$, requires assumptions on the volume of LLSs and DLAs, which are model-dependent. An example of such model, and the inferred value of $\langle f_{\rm HI} \rangle_V$, is discussed in § 2.2.3.

2.2.3 Theoretical expectations for H I

The goal of this section is to highlight the need for a hybrid web-bubble model to interpret high-z galaxy observations. We present theoretical estimates of $\langle f_{\rm HI}\rangle_V$ using analytic models for the three different classes of ionization structure in the IGM. These calculations illustrate the redshift evolution of $\langle f_{\rm HI}\rangle_V$ in the web (§ 2.2.3), bubble (§ 2.2.3), and hybrid web-bubble (§ 2.2.3) model.

H I Fraction in the Web Model

In the web model, $\langle f_{\rm HI}\rangle_V$ is expected to increase with increasing redshift due to decreasing photoionization rate, and/or increasing mean gas density (by Hubble expansion). $\langle f_{\rm HI}\rangle_V$ can be estimated as (e.g [309, 43])

$$\langle f_{\rm HI}\rangle_V = \int_0^{\Delta_{ss}} x_{\rm HI}(\Delta_b)P(\Delta_b)d\Delta_b + \int_{\Delta_{ss}}^{\infty} P(\Delta_b)d\Delta_b, \tag{2.1}$$

where Δ_b is the baryon overdensity, $P(\Delta_b)$ is the volume-weighted overdensity probability distribution function[13], and $\Delta_{ss} \propto \Gamma^{2/3}$ is the density threshold above which the gas self-shields against the UV background ([388, 158]). $x_{\rm HI}(\Delta_b) = \alpha_A(T)\bar{n}_H^{com}(1+z)^3 f_e\Delta_b/\Gamma$ is the neutral fraction obtained assuming local photoionization equilibrium with a uniform photoionization rate Γ ($\rm s^{-1}$), $\bar{n}_H^{com}$ is the average comoving hydrogen number density, α_A is the case A recombination rate at temperature T, and f_e is the electron fraction per hydrogen atom. The first and second term on the right hand side of equation (2.1) are the volume-weighted H I fraction embedded in residual H I in the diffuse IGM and the self-shielded gas, respectively.

We consider two models for the redshift evolution of the photoionization rate Γ: the CONST model assumes a constant $\Gamma = \Gamma(z = 4.75)$, while the DEC model assumes a photoionization rate decreasing with increasing redshift, i.e. $\Gamma(z) = \Gamma(z = 4.75)[(1 + z)/5.75]^{-1.5}$ ([63])[14].

The blue and cyan lines in Fig. 2.2 show an example of the redshift evolution of $\langle f_{\rm HI}\rangle_V$, with the two different contributions from residual H I in the diffuse IGM (residual H I; cyan) and neutral self-shielded gas (self-shielded H I; blue). The global H I fraction is clearly dominated at all redshifts by the self-shielded gas. While, as expected, the H I fraction increases with redshift due to the larger mean gas density, $\langle f_{\rm HI}\rangle_V$ increases more markedly in the DEC model due to the lower photoionization rate.

[13] $P(\Delta_b)$ is adopted from [309] with the redshift extrapolation of [12].

[14] $\Gamma(z = 4.75)$ is chosen to be consistent with the UV background measurement from the Lyα forest, i.e. $\log_{10}\Gamma(z = 4.75)/(10^{-12}\rm s^{-1}) = -0.029^{+0.156}_{-0.147}$ ([18]).

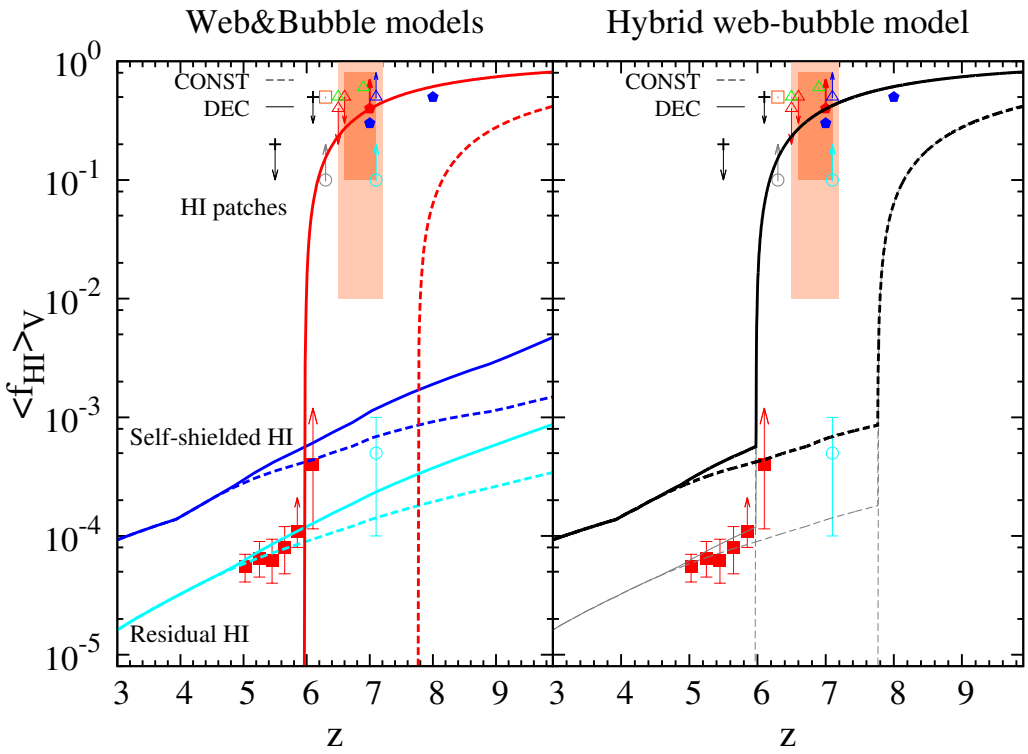

Figure 2.2: Redshift evolution of the volume-weighted H I fraction obtained with the analytic models. The blue and cyan lines refer to the residual neutral gas in the ionized regions (residual H I, cyan) and self-shielded gas (self-shielded H I, blue), i.e. the two different reservoirs in the web model. The red lines refer to the bubble model, where patchy neutral gas outside ionized bubbles (H I patches) is the main reservoir of H I. Results for the web-bubble model are shown in the right panel for a case in which the residual H I inside ionized bubbles is calculated including only the diffuse (gray lines) or the self-shielded (black lines) gas. CONST and DEC models are indicated as dashed and solid lines. The points in Fig. 2.1, inferred from observations, are overlaid as a guide, while the faint and dark orange boxes refer to our estimate based on the Lyα luminosity function alone and used in combination with the equivalent width distribution, respectively. This plot illustrates how hybrid models are required to explain the observational constraints at all redshifts.

H I Fraction in the Bubble Model

In the bubble model, the time evolution of the volume filling factor of ionized bubbles, the 'porosity' factor Q_i, is given by (e.g. [261])

$$\frac{dQ_i(t)}{dt} = \frac{\dot{n}_{ion}(t)}{\bar{n}_H^{com}} - \frac{Q_i(t)}{\bar{t}_{rec,B}(t)}, \tag{2.2}$$

where $\dot{n}_{ion}$ is the ionizing photon emissivity (in units of ph s^{-1} cm^{-3}) and $\bar{t}_{rec,B} = [\alpha_B \bar{n}_H^{com}(1+z)^3 C]^{-1}$ is the case B recombination timescale with clumping factor C. If 'empty' bubbles are assumed, we have

$$\langle f_{\rm HI} \rangle_V = 1 - Q_i. \tag{2.3}$$

As with the web model, we consider two cases: the CONST model assumes a constant ionizing photon emissivity[15] $\dot{n}_{ion}(z) = \dot{n}_{ion}(z = 4.75)$, while the DEC model assumes that the ionizing photon emissivity decreases with increasing redshift, i.e. $\dot{n}_{ion}(z) = \dot{n}_{ion}(z = 4.75)[(1+z)/5.75]^{-1.5}$. This choice of redshift evolution is made to bracket the possible range of parameters satisfying the Gunn-Peterson constraints ([140]).

In the left panel of Fig. 2.2 the redshift evolution of $\langle f_{\rm HI} \rangle_V$ in the bubble model (red lines) shows a rapid change at $z \sim 6-8$, when $\langle f_{\rm HI} \rangle_V$ plummets to zero once reionization ends. A smooth transition to the post-reionized IGM, where small-scale absorbers must exist, is clearly absent from these models as no H I gas is present inside ionized bubbles. The behaviour in the CONST and DEC cases is very similar, with an earlier reionization in the former case, where a larger photoionization rate is present.

H I Fraction in the Web-Bubble Model

In the web-bubble model we assume that (*i*) gas inside ionized bubbles behaves as in the web model, and (*ii*) gas outside ionized bubbles is fully neutral. These assumptions lead to

$$\langle f_{\rm HI} \rangle_V = 1 - \left[1 - \int_0^{\Delta_{ss}} x_{\rm HI}(\Delta_b) P(\Delta_b) d\Delta_b - \int_{\Delta_{ss}}^{\infty} P(\Delta_b) d\Delta_b \right] Q_i, \tag{2.4}$$

where the terms in square brackets are the H II fraction inside the ionized bubbles.

The redshift evolution of $\langle f_{\rm HI} \rangle_V$ in the web-bubble model is shown in the right panel of Fig. 2.2 for a case in which the residual H I inside ionized bubbles is calculated including only the diffuse (gray lines) or the self-shielded (black lines) gas. The assumed values of the photoionization rate and ionizing photon emissivity are the same as used in the previous sections. A comparison between the left and right panels of the figure shows that the web-bubble model produces a smooth transition from the bubble model (patchy reionization) during the EoR to the web model (dominated by small-scale absorbers) in post-reionization.

[15] $\dot{n}_{ion}(z)$ is anchored at $z = 4.75$ based on Lyα forest constraints, indicating $-0.46 \lesssim \log_{10} \dot{n}_{ion}/10^{51}$ph s^{-1}cMpc$^{-3} \lesssim 0.35$ over $2.40 \lesssim z \lesssim 4.75$ ([18]). Here we assume $\dot{n}_{ion}(z = 4.75) = 10^{51}$ph s^{-1}cMpc^{-3}.

Hence, to coherently explain and interpret present observations, a unified framework that includes both large-scale bubbles *and* small-scale absorbers is essential because (*i*) the presence of small-scale absorbers at lower-z is evident from observations (Fig. 2.1), and (*ii*) a smooth transition from a patchy reionization to a post-reionized IGM with small-scale absorbers is only possible within a hybrid web-bubble model (Fig. 2.2).

2.3 Cosmological Lyα radiative transfer

In this section we present the formalism adopted to follow the cosmological Lyα transfer through the reionization models discussed above.

The general equation describing line transfer in the Lagrangian fluid frame is ([306, 68, 295, 116])

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \boldsymbol{n}\cdot\boldsymbol{\nabla} I_\nu - \frac{H + \boldsymbol{n}\cdot\boldsymbol{\nabla}\boldsymbol{v}\cdot\boldsymbol{n}}{c}\nu\frac{\partial I_\nu}{\partial \nu} + 3\frac{H}{c}I_\nu = -\sigma_\alpha n_{\rm HI}\varphi_\nu I_\nu + \sigma_\alpha n_{\rm HI}\int \mathcal{R}(\nu,\nu')J_\nu(\nu')d\nu' + \varepsilon_\nu, \tag{2.5}$$

where I_ν is the specific intensity, J_ν is the angle-averaged intensity, ε_ν is the Lyα emissivity, $\boldsymbol{v}$ is the peculiar velocity, $\boldsymbol{n}$ is the unit direction vector of rays, $\sigma_\alpha = 0.011$ cm^{2}Hz is the Lyα cross section, and φ_ν is the line profile of the Lyα resonance line (units Hz^{-1}). The $\boldsymbol{n}\cdot\boldsymbol{\nabla}\boldsymbol{v}\cdot\boldsymbol{n}$ term is the Doppler shift effect and $\mathcal{R}(\nu,\nu')$ is the redistribution function describing the resonant scattering of Lyα photons.

There are generally no analytic solutions to equation (2.5). However, by performing a separation of scales, the problem can be simplified: multiple scattering effects are predominant on ISM scales because the surface brightness of Lyα photons that are scattered back into the line-of-sight at IGM scales is typically negligibly small. As scatterings on such small ISM scales can be effectively treated as a modification of the intrinsic line profile, and the scattering term can be overall neglected ([245]).

Equation (2.5) can then be readily integrated along a line-of-sight to give the so-called '$e^{-\tau}$ approximation' (e.g. [183, 295, 284]). In this approximation, the Lyα flux F_α observed from a Lyα-emitting galaxy at redshift z_s is given by

$$F_\alpha = \frac{L_\alpha}{4\pi D_L^2(z_s)}\int S_\nu(\nu_e)e^{-\tau_\alpha(\nu_e)}d\nu_e = \frac{L_\alpha \mathcal{T}_{IGM}}{4\pi D_L^2(z_s)}, \tag{2.6}$$

where ν_e is the frequency of the Lyα photon when it is emitted, $D_L(z_s)$ is the luminosity distance, L_α is the intrinsic bolometric Lyα luminosity (in units of erg s^{-1}), $S_\nu(\nu_e)$ (in units of Hz^{-1}) is the effective intrinsic line profile (including the effect of the ISM/CGM) normalized such that $\int S_\nu(\nu_e)d\nu_e = 1$. $\mathcal{T}_{IGM} = \int S_\nu(\nu_e)e^{-\tau_\alpha(\nu_e)}d\nu_e$ denotes the IGM transmission factor ([121]), and the Lyα optical depth $\tau_\alpha(\nu_e)$ is

$$\tau_\alpha(\nu_e) \approx \sigma_\alpha \int_0^{l_p} dl'_p n_{\rm HI}(l'_p)\varphi_\nu\left[T, \nu_e\left(1 - \frac{v_{tot}(l'_p)}{c}\right)\right], \tag{2.7}$$

where T is the gas temperature and $v_{tot} = H(z_s)l_p + v(l_p)$ is the sum of the Hubble flow and the peculiar velocity. It is customary to express ν_e in terms of a velocity shift, i.e. $\Delta v/c = 1 - \nu_e/\nu_\alpha$. In the following we will use this convention.

We would like to note here that by using the $e^{-\tau}$ approximation we ignore photons that scatter back into the line-of-sight, which would give rise to a low surface brightness 'fuzz'. [245] compared the $e^{-\tau}$ approximation to a full Monte-Carlo Lyα radiative transfer approach finding that the $e^{-\tau}$ approximation provides a good description of the transfer through the IGM as long as this is assumed to start at a distance larger than 1.5 times the virial radius of the dark matter halo hosting a Lyα galaxy. We have verified that this condition is met throughout our work.

We also introduce the mean IGM transmission factor and effective optical depth to characterize the typical impact of the intergalactic environment around Lyα-emitting galaxies. The mean Lyα flux of many Lyα-emitting galaxies is $\langle F_\alpha \rangle \approx \langle L_\alpha \rangle \langle \mathcal{T}_{IGM} \rangle / (4\pi D_L^2)$, where

$$\langle \mathcal{T}_{IGM} \rangle \approx \int \langle S_\nu(\nu_e) \rangle e^{-\tau_\alpha^{\rm eff}(\nu_e)} d\nu_e \tag{2.8}$$

is the mean IGM transmission factor and $\tau_\alpha^{\rm eff} = -\ln\langle e^{-\tau_\alpha} \rangle$ is the effective optical depth (e.g. [183]). Here we have assumed that the intrinsic line profiles of Lyα galaxies and the optical depth of the IGM are uncorrelated, i.e. that $\langle S_\nu(\nu_e) e^{-\tau_\alpha(\nu_e)} \rangle \approx \langle S_\nu(\nu_e) \rangle e^{-\tau_\alpha^{\rm eff}(\nu_e)}$.

The optical depth contribution from different intervening IGM absorbers (the diffuse neutral IGM outside ionized bubbles and the small-scale absorbers) is additive, i.e. $\tau_\alpha = \tau_{bub} + \tau_{web}$. The same applies to the effective optical depth, i.e. $\tau_\alpha^{\rm eff} = \tau_{bub}^{\rm eff} + \tau_{web}^{\rm eff}$. In the bubble [web] model of reionization we ignore τ_{web} [τ_{bub}], while in the web-bubble model we include both. These two terms are discussed in more detail in the following sections.

2.3.1 Lyα Opacity from Large-Scale H I Patches

In the bubble model the Lyα optical depth is due to diffuse expanding neutral IGM outside ionized bubbles (H I patches). The Lyα optical depth of a homogeneous H I patch extending between comoving distance from a Lyα-emitting galaxy R_1 and R_2, can be written as ([302, 116])

$$\tau_{patch}(\nu_e, R_1, R_2) = \tau_{GP} \int_{x(\nu_e, R_2)}^{x(\nu_e, R_1)} \phi(x) dx, \tag{2.9}$$

where $\phi = \Delta\nu_D \varphi_\nu$ is the dimensionless line profile, $\tau_{GP} = c\sigma_\alpha \bar{n}_{\rm H}^{com}(1+z_s)^3/(\nu_\alpha H) \approx 4.44 \times 10^5 [(1+z_s)/7.6]^{3/2}$ is the Gunn-Peterson optical depth, and $x(\nu_e, R) = \{\nu_e(1 - H(z_s)R/[(1+z_s)c]) - \nu_\alpha\}/\Delta\nu_D$. $\Delta\nu_D = \frac{\nu_\alpha}{c}\sqrt{\frac{2k_BT}{m_p}}$ is the Doppler width, with k_B Boltzmann constant and m_p proton mass.

In general, the Lyα optical depth along a line-of-sight in the bubble model is given by

$$\tau_{bub}(\nu_e) = \sum_{i=1} \tau_{patch}(\nu_e, R_{1,i}, R_{2,i}), \tag{2.10}$$

where $R_{1,i}$ [$R_{2,i}$] is the near [far] side of the edge of the i-th H I patch. The effective optical depth through an ensemble of H I patches is

$$e^{-\tau_{bub}^{\rm eff}(\nu_e)} = \int e^{-\tau_{bub}(\nu_e)} P[\tau_{bub}(\nu_e)] d\tau_{bub}(\nu_e), \tag{2.11}$$

where $P[\tau_{bub}(\nu_e)]$ denotes the probability distribution for $\tau_{bub}(\nu_e)$, which must be obtained from cosmological realizations of the bubble model.

There is a simpler limiting analytic case if we assume an ensemble of single large H I patches along all lines-of-sight. In the limit of a large H I patch ($R_2 \rightarrow \infty$), the optical depth along a line-of-sight is $\tau_{patch}(\nu_e, R_1) \approx \frac{\tau_{GP}\Lambda}{4\pi^2\nu_\alpha} \left| \frac{\nu_e}{\nu_\alpha} \left[1 - \frac{H(z_s)R_1}{(1+z_s)c} \right] - 1 \right|^{-1}$, where $\Lambda = 6.25 \times 10^8$ s^{-1} is the damping coefficient (e.g. [307, 285]). Then, for an ensemble of large H I patches we can evaluate the effective optical depth as

$$e^{-\tau_{bub}^{\rm eff}(\nu_e)} \approx \int e^{-\tau_{patch}(\nu_e, R_1)} P(R_1) dR_1, \tag{2.12}$$

where $P(R_1)dR_1$ is the probability to find the near side of a H I patch at a distance R_1 from a Lyα-emitting galaxy (for a related definition of bubble size distribution, see [301]). We model $P(R_1)$ as a Schechter function, $P(R_1) \propto R^{\alpha_1} \exp(-R_1/R_*)$, normalized as $\int P(R_1)dR_1 = 1$; α_1 and R_* are free parameters. We compare this analytic estimate of the effective optical depth to numerical calculations in § 2.5.2.

2.3.2 Lyα Opacity from Small-Scale Absorbers

In the web model the H I gas is distributed in a collection of self-shielded absorbers. Each absorber is characterized by its H I column density, $N_{\rm HI}$, and its proper velocity, v_c, relative to a given Lyα-emitting galaxy. The Lyα optical depth through a single absorber is

$$\tau_{abs}(\nu_e) = \sigma_\alpha N_{\rm HI} \varphi_\nu \left[T_c, \nu_e \left(1 - \frac{v_c}{c} \right) \right], \tag{2.13}$$

where T_c denotes the gas temperature of an absorber.

We introduce a novel analytic model of the Lyα opacity from small-scale absorbers as follows. The effective optical depth of an ensemble of H I absorbers surrounding a Lyα-emitting galaxy is (see Appendix A.2 for a derivation)

$$\tau_{web}^{\rm eff}(\nu_e) = \int dN_{\rm HI} \frac{\partial^2 \mathcal{N}}{\partial N_{\rm HI} \partial z} \left| \frac{dz}{dl_p} \right| \int \frac{dv_c}{H(z_s)} \left[1 + \xi_v(v_c, N_{\rm HI}) \right] \left[1 - e^{-\tau_{abs}(\nu_e, v_c, N_{\rm HI})} \right], \tag{2.14}$$

where $\xi_v(v_c, N_{\rm HI})$ is the galaxy-absorber correlation function in velocity space. We refer to a Gaussian streaming model (GSM) for $\xi_v(v_c, N_{\rm HI})$ when

$$1 + \xi_v(v_c, N_{\rm HI}) = \int \frac{aH dr_{12}}{\sqrt{2\pi\sigma_{12}^2(r_{12})}} \left(1 + \xi(r_{12}) \right) \exp \left[-\frac{(v_c - aHr_{12} - \langle v_{12}(r_{12}) \rangle)^2}{2\sigma_{12}^2(r_{12})} \right], \tag{2.15}$$

where r_{12} is the comoving separation between a galaxy and an absorber, $\xi(r_{12})$ is the real-space galaxy-absorber correlation function, $\langle v_{12}(r_{12}) \rangle$ [$\sigma_{12}(r_{12})$] is the mean radial pairwise velocity [the pairwise velocity dispersion] between galaxy-absorber pairs, and $a = (1+z_s)^{-1}$ is the scale factor.

The Region of Influence

As the optical depth depends on v_c, it is useful to calculate the 'critical' velocity, $v_{\rm crit}$, at which the optical depth of an absorber to a Lyα photon emitted at frequency ν_e becomes unity for a given H I column density, i.e. $\tau_{abs}(v_c = v_{\rm crit}) = 1$. In fact, to first order, the Lyα visibility is only affected by small-scale absorbers moving away from a central Lyα-emitting galaxy with $v_c < v_{\rm crit}$. We refer to the region that contains these absorbers as the 'region of influence'. For high-column density absorbers such as LLS/DLA, the above condition is met in the wing of the absorption line profile $\varphi_\nu \approx \Lambda\,[4\pi^2(\nu_e(1 - v_c/c) - \nu_\alpha)^2]^{-1}$. From the Lorentz wing it follows that for an absorber with H I column density $N_{\rm HI}$,

$$\frac{v_{\rm crit}}{c} = 1 - \frac{\nu_\alpha}{\nu_e}\left(1 - \sqrt{\frac{\sigma_\alpha N_{\rm HI}\Lambda}{4\pi^2\nu_\alpha^2}}\right). \tag{2.16}$$

If we set $\nu_e = \nu_\alpha$, then $v_{\rm crit} = c\sqrt{\frac{\sigma_\alpha N_{\rm HI}\Lambda}{4\pi^2\nu_\alpha^2}} = 507.3(N_{\rm HI}/10^{20}{\rm cm}^{-2})^{1/2}$ km s^{-1}. For a pure Hubble flow, the critical velocity corresponds to the comoving distance

$$D_{\rm infl} = \frac{v_{\rm crit}}{H_0}\frac{1+z}{[\Omega_m(1+z)^3 + \Omega_\Lambda]^{1/2}}. \tag{2.17}$$

As a reference, $D_{\rm infl} = 3.5(N_{\rm HI}/10^{20}{\rm cm}^{-2})^{1/2}h^{-1}$cMpc at $z = 7$.

Armed with the analytic framework of Lyα transfer described above to aid the understanding of our results, in the next section we perform cosmological hydrodynamical, radiative transfer simulations and derive a mock survey of Lyα-emitting galaxies with various reionization models.

2.4 Simulations

In this section we describe the simulations used to model the observability of high redshift Lyα-emitting galaxies, and the mock galaxy catalogue obtained from them.

2.4.1 Hydrodynamical Simulations of the IGM

We employ the AMR cosmological N-body/hydrodynamical code RAMSES ([439]) to simulate the IGM in a box of size $25h^{-1}$cMpc on a side. The cosmological parameters are $(\Omega_m, \Omega_\Lambda, \Omega_b, \sigma_8, n_s, h) = (0.26, 0.74, 0.044, 0.85, 0.95, 0.72)$. The total number of dark matter particles is 256^3 with mass resolution of $m_{\rm DM} = 5.5 \times 10^7 h^{-1}{\rm M}_\odot$, and the gas is included on the 256^3 base grid ($97.7h^{-1}$ckpc cell size) with two levels of refinement, reaching a 1024^3 grid ($24.4h^{-1}$ckpc cell size) at the highest AMR refinement level. For our choice of the simulation parameters, the cosmological Jeans length is $\sim 57(T/100{\rm K})^{1/2}h^{-1}$ckpc, which corresponds to a Jeans mass of $\sim 1.3 \times 10^7 h^{-1}{\rm M}_\odot$ at $T = 100$ K. The finest spatial resolution therefore fulfills the minimum requirement to resolve self-shielded absorbers of order of the Jeans length ([388]).

Table 2.1: List of reionization models. The columns indicate, from left to right, the model and its name, the photoionization rate in terms of $\Gamma_{-12} = \Gamma/10^{-12}$ s^{-1} as assumed at $z = 6$ [7] in the bubble [web] model, and the resulting volume-weighted H I fraction, $\langle f_{\rm HI}\rangle_V$, at $z = 7$.

bubble model	$\Gamma_{-12}(z{=}6)$	$\langle f_{\rm HI}\rangle_V(z{=}7)$
B1	0.380	0.365
B2	0.190	0.676
B3	0.019	0.990
web model	$\Gamma_{-12}(z{=}7)$	$\langle f_{\rm HI}\rangle_V(z{=}7)$
W1	0.1	0.00073
W2	0.01	0.012
W3	0.005	0.032
web-bubble model		$\langle f_{\rm HI}\rangle_V(z{=}7)$
B1+W2		0.373
B1+W3		0.387
B2+W2		0.681
B2+W3		0.688

The initial conditions are generated with the COSMICS package ([26]). The initial temperature is set to 650 K. This is higher than the value expected from the cooling and heating of the IGM after recombination as computed from RECFAST ([396]) to take into account the fact that our calculation did not include primordial heating mechanisms such as Compton heating by the CMB. The temperature is then calculated according to an adiabatic evolution. The initial redshift of the simulation is chosen as $z_{ini} = 225$, to allow sufficient nonlinearity to develop at the reionization epoch $z \sim 20 - 5$.

The dark matter haloes are identified using the HOP algorithm ([134]) as implemented in the RAMSES package.

2.4.2 Radiative Transfer Simulations

We use a two-way approach to follow the radiative transfer. In the first, we generate bubble models by performing full radiative transfer simulations of ionizing UV photons by post-processing the base AMR grid of the cosmological hydrodynamical simulation (§ 2.4.2). In the second, we generate web models by post-processing the finest AMR grid without full RT, but assuming photoionization equilibrium with a uniform UV background and a self-shielding criterion (§ 2.4.2). To generate the web-bubble models we modify the bubble models by treating the regions inside the ionized bubbles as web models (§ 2.4.2). This approach enables us to access spatial scales for the self-shielding gas which are presently beyond the computational capability of the full radiative transfer simulations.

We emphasize that the purpose of these simulations is not to produce the best possible reionization model, but to explore the impact of large-scale patchy reionization features (i.e.

ionized bubbles) and small-scale absorbers on the observability of Lyα-emitting galaxies and on the inference of $\langle f_{\rm HI}\rangle_V$ using Lyα surveys.

Bubble Models

We use the cosmological radiative transfer code CRASH ([81, 270, 269, 177]) to generate our bubble models. For the full RT computation, we post-process the density and temperature fields on the 256^3 cells of the hydrodynamical simulation. While our box size is not sufficient to include the largest ionized bubbles present during the later stages of reionization, this does not affect the goal of the chapter.

The model for the ionizing sources is based on the one described in [80]: the volume averaged ionizing emissivity, $\dot{n}_{ion}$ (photons s^{-1} cMpc^{-3}), at $z > 6$ is parameterized as $\dot{n}_{ion}(z) = 10^{50.89}\chi(z)\frac{\alpha_b+3}{2\alpha}\left(\frac{\Gamma_{-12}(z=6)}{0.19}\right)$, where $\chi(z) = ae^{b(z-9)}\left[a-b+be^{a(z-9)}\right]^{-1}$, with $a = 14/15$ and $b = 2/3$ (see [43]). The values of $\Gamma_{-12}(z=6)$ are shown in Table 2.1. We assume that the ionizing emissivity is produced by galaxies with a power-law spectrum of slope $\alpha_b = \alpha = 3$, and we distribute it among all haloes proportionally to their mass.

We ran the radiative transfer simulation using 10 gas density and temperature snapshots from $z = 15$ to $z = 5$ equally spaced in redshift, including both hydrogen and helium with a number fraction 0.92 and 0.08, respectively. For each source, we emit 10^6 photon packets distributed according to the power-law spectrum with 29 frequency bins sampled from 13.6eV to 200eV.

Finally, we produce a catalogue of bubble models for different values of $\Gamma_{-12}(z=6)$. Slices through these models are shown in the top panels of Figs. 2.3 and 2.4. In Fig. 2.3 the maps of H I number density clearly show that the global H I fraction increases as $\Gamma_{-12}(z=6)$ decreases (from left to right), as expected. More specifically, a volume-weighted H I fraction of $\langle f_{\rm HI}\rangle_V = 0.365$, 0.676 and 0.990 is obtained at $z = 7$ for $\Gamma_{-12}(z=6) = 0.380$, 0.190 and 0.019, respectively. Furthermore, the figures show the characteristic feature of patchy reionization, i.e. large-scale bubbles.

Since there is a one-to-one correspondence between $\Gamma_{-12}(z=6)$ and $\langle f_{\rm HI}\rangle_V(z=7)$, we will use them interchangeably to specify the model.

Web Models

We use the prescription of Rahmati et al. ([359]) to account for self-shielding gas in the web models, which consists of a fitting function matched to their full RT transfer simulation. This prescription assumes photoionization equilibrium in each cell of the simulation with a modified background (see below), i.e.

$$\alpha_A(T)n_{\rm H\,II}n_e = \Gamma_{Rah}n_{\rm HI}, \tag{2.18}$$

where n_e is the electron number density, and Γ_{Rah} is the modified local photoionization rate. The neutral fraction in each cell is then given by $x_{\rm HI} = \left[(\gamma+2) - \sqrt{(\gamma+2)^2-4}\right]/2$, where $\gamma = \Gamma_{Rah}/(\alpha_A n_{\rm H} f_e)$. The factor $f_e = n_e/n_{\rm H\,II}$ is 1 for a pure hydrogen medium, while

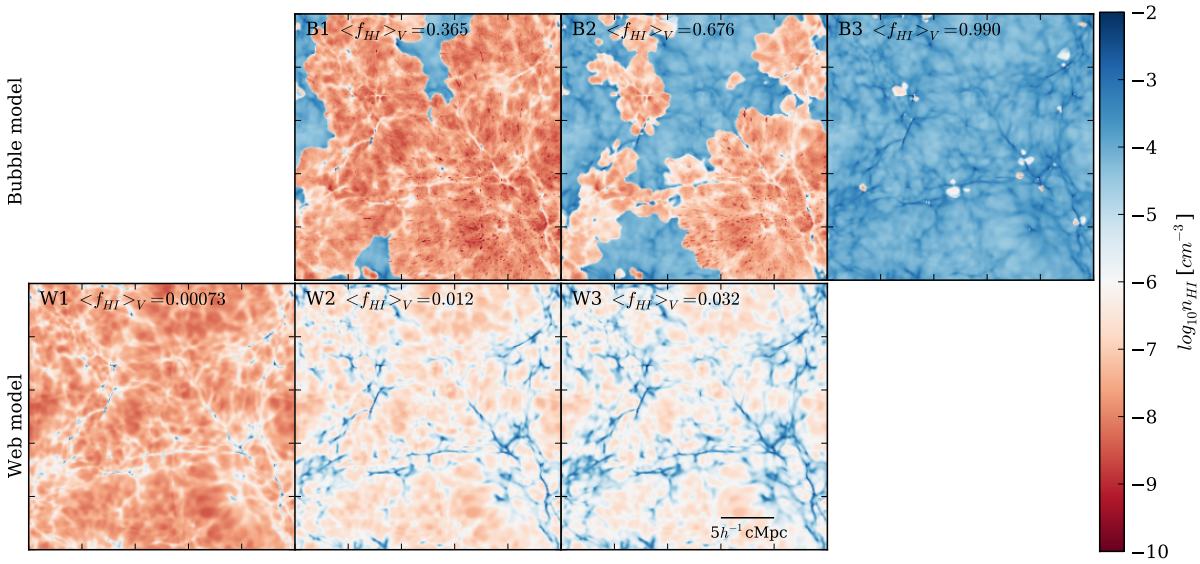

Figure 2.3: Neutral hydrogen number density, $n_{\rm HI}$, at $z = 7$ in slices of our simulations for the bubble (B1, B2 and B3; *top panels*) and web (W1, W2 and W3; *bottom panels*) models detailed in Table 2.1. Each snapshot is a x-y slice at $12.5h^{-1}$cMpc with $97.7h^{-1}$ckpc thickness. Panels in the same column give a similar suppression of the Lyα visibility in the observed Lyα luminosity functions shown in Fig. 2.6.

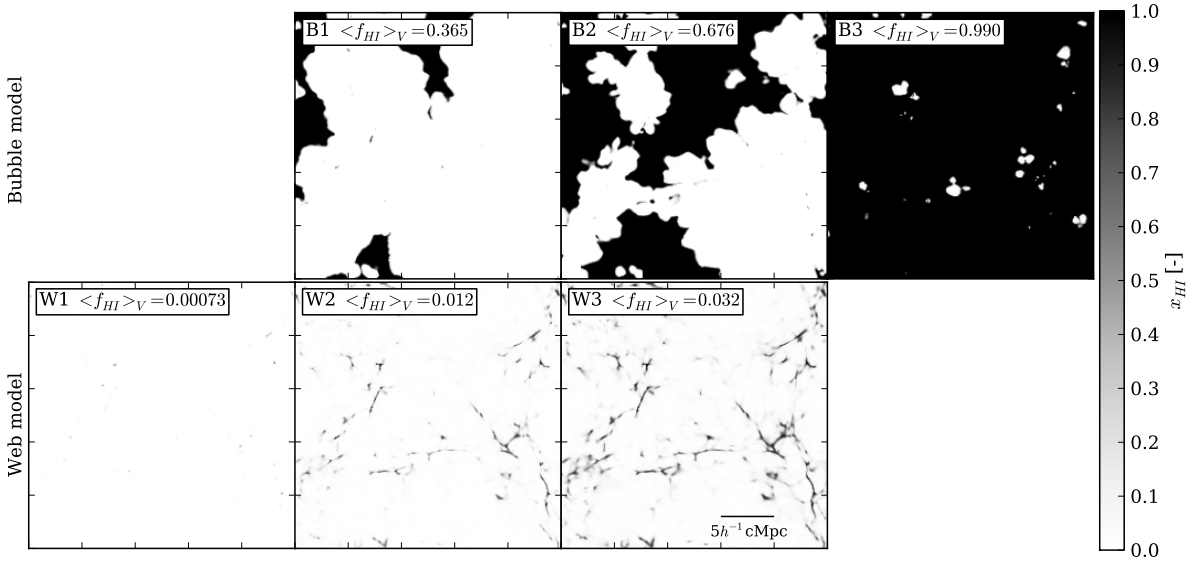

Figure 2.4: Same as in Fig. 2.3 but for the local H I fraction $x_{\rm HI}$.

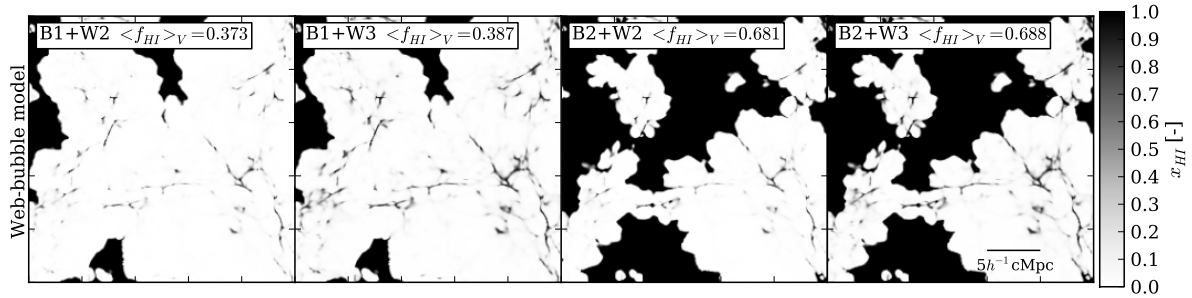

Figure 2.5: Slice map of the local H I fraction for hybrid web-bubble models.

$f_e > 1$ if helium is included. We assume that the IGM temperature is $T \sim 10^4$K due to photoionization heating. The modified local photoionization rate is given by

$$\frac{\Gamma_{Rah}}{\Gamma} = 0.98\left[1+\left(\frac{n_{\rm H}}{n_{SS}}\right)^{1.64}\right]^{-2.28} + 0.02\left[1+\left(\frac{n_{\rm H}}{n_{SS}}\right)\right]^{-0.84}, \tag{2.19}$$

where n_{SS} is the density at which the gas starts to be self-shielded

$$n_{SS} = 6.73\times10^{-3}\Gamma_{-12}^{2/3}\left(\frac{T}{10^4{\rm K}}\right)^{0.17}{\rm cm}^{-3}. \tag{2.20}$$

To compute $n_{\rm HI}$ using the above prescription we use the density field of the finest AMR level 1024^3 from the hydrodynamical simulation[16].

The values adopted for the uniform photoionization rate Γ_{-12} are found in Table 2.1. Slices through our web models are reported in the lower panels of Figs. 2.3 and 2.4. Similarly to the bubble models, the maps show a higher neutral fraction for decreasing photoionization rate. However, the ionization structure of the IGM is significantly different, as the neutral gas is concentrated in high density peaks where small-scale absorbers, whose distribution follows the structure of the cosmic web, reside[17].

Web-Bubble Models

We generate the web-bubble models at $z = 7$ as follows. First, we take a full RT simulation used to generate the bubble models (B1 and B2). Then, we recalculate the local H I

[16]As a comparison, we have also calculated $n_{\rm HI}$ using a threshold method, in which all the cells with gas density above n_{SS} are assumed to be fully neutral, otherwise the neutral fraction is computed assuming photoionization equilibrium with Γ rather than Γ_{Rah}, i.e. $\alpha_A(T)n_{\rm H\,II}n_e = \Gamma n_{\rm HI}$. We note that while mapping between assumed photoionization rate and the abundance of small-scale absorbers changes depending on the prescription, as long as $\langle f_{\rm HI}\rangle_V$ embedded inside small-scale absorbers is similar, the result is insensitive to the self-shielding prescription. Hence, the quantity that more directly impact the observation of Lyα-emitting galaxies is the number density of small-scale absorbers rather than the photoionization rate.

[17]Note that web models are *not* equivalent to outside-in reionization scenarios. They simply show the region of the universe that is reionized early in an inside-out scenario, with residual self-shielded H I.

fraction inside the ionized bubbles according to the web model with a photoionization rate $\Gamma_{-12}(z = 7) = 0.01$ and 0.005 (W2 and W3) on the finest AMR grid[18]. In practice, $x_{\rm HI}$ is calculated locally as the maximum between the values obtained from the bubble and the web model. Our web-bubble models are catalogued in Table 2.1.

Slices through the web-bubble models are shown in Fig. 2.5 in terms of $x_{\rm HI}$ map. As expected, the evolution of $x_{\rm HI}$ with photoionization rate is the same as the one in the web and bubble models. Quantitatively, though, the neutral fraction here is slightly higher than the one in the corresponding bubble models due to the contribution of small-scale absorbers (see Table 2.1). In addition, the ionization structure of the IGM looks like a combination of the one from the bubble and web models, as the small-scale absorbers appear as impurities inside large-scale ionized bubbles.

2.4.3 Mock Galaxy Catalogue

The observed Lyα luminosity of a galaxy is related to its intrinsic Lyα luminosity via the IGM transmission factor $\mathcal{T}_{IGM}$ as $L_\alpha^{obs} = \mathcal{T}_{IGM} L_\alpha$ (this is discussed more in detail in § 2.3; see also Appendix A.4 for more technical aspects). We stress that 'intrinsic' here refers to the Lyα luminosity that a galaxy would have if the IGM were transparent. As our main results are insensitive to the precise model for the intrinsic luminosity function (LF), we only briefly describe the methodology applied to generate the intrinsic mock galaxy catalogue.

We use the abundance matching technique (e.g. [342]; see also Appendix A.3) to populate dark matter haloes with Lyα emitting galaxies. We find the relation between halo mass and intrinsic Lyα luminosity by equating the *observed* cumulative Lyα luminosity function $n(> L_\alpha)$ (in units of cMpc^{-3}) at $z = 5.7$ ([324]) to the simulated halo mass function $dn(> M_h')/dM_h'$ at $z = 7$,

$$n(> L_\alpha) = f_{duty} \int_{M_h(L_\alpha)}^{\infty} \frac{dn(> M_h')}{dM_h'} dM_h', \tag{2.21}$$

where f_{duty} is the duty cycle and $M_h(L_\alpha)$ is the halo mass corresponding to a Lyα luminosity L_α. We thus assume that the intrinsic Lyα luminosity function at $z = 7$ is equal to the observed one at $z = 5.7$, and that the difference between $z = 5.7$ and $z = 7$ is entirely due to the IGM. We therefore constrain the IGM opacity using the *variation* of the Lyα LF relative to that in the post-reionization Universe [19]. The abundance matching technique gives a semi-empirical relation between the halo mass and the intrinsic Lyα luminosity for

[18] While in principle the photoionization rate inside bubbles is not independent of bubble size, we take this as a convenient free *ad hoc* parameter to adjust the abundance of self-shielded absorbers inside bubbles.

[19] The small-scale absorbers in the post-reionized universe may affect the visibility of Lyα-emitting galaxies at $z < 6$. Hence, calibrating with $z < 6$ Lyα LF may not give a truly 'intrinsic' Lyα luminosity as defined above. This contribution should in principle be subtracted. However, as we will confirm in § 2.5.3, the impact of small-scale absorbers at $z \lesssim 6$ is small. Note, though, that ignoring the post-reionization optical depth of the IGM causes us to *underestimate* the intrinsic Lyα luminosity of galaxies, which in turn leads us to underestimate the IGM opacity and hence the neutral fraction in the IGM (see [121]).

each f_{duty} (examples are shown in Fig. A.2). In our fiducial case we use $f_{duty} = 1$. We then populate each halo with a single Lyα-emitting galaxy of intrinsic Lyα luminosity given by the $M_h - L_\alpha$ relation.

Because observations are available only down to $\log_{10}[L_\alpha/(\mathrm{erg\,s^{-1}})] \approx 42.5$, to extend the calculations to lower luminosities we extrapolate assuming a faint-end slope of 1.5 ([324], but see [181, 130] for both theoretical and observational support for significantly steeper slopes of ≈ 2.2). We note that, because of the small box size (which is needed to include small-scale absorbers and the faint galaxies responsible for reionization), the simulated LFs only extend to $\log_{10}[L_\alpha/(\mathrm{erg\,s^{-1}})] \approx 42.8$.

We model the Lyα transfer in the ISM/CGM through the Lyα spectral line profile (e.g. [121, 213, 77]), by assuming a Gaussian profile with circular velocity $v_{circ} = 20.4h^{1/3}(M_h/10^8\mathrm{M}_\odot)^{1/3}[(1+z)/7.6]^{1/2}$ km s^{-1} ([384, 501]), shifted redward by $\Delta v = 600$ km s^{-1} to mimic the effect of scattering through a galactic wind ([122, 121]). This is rather arbitrary, but [428] and [479] justify a number between 200-800 km s^{-1}. While the quantitative results are affected by this choice, the qualitative conclusions in this chapter remain valid. We point out that our model assumes a universal line profile and shift, while a distribution is more likely. Since $\mathcal{T}_{IGM}$ is highest for redshifted Lyα lines, this can bias samples of Lyα-selected galaxies to larger Δv.

2.5 Results

2.5.1 Lyα Luminosity Function

We first show the impact of large-scale H I patches and small-scale absorbers on the Lyα luminosity function in Fig. 2.6, which contains the differential intrinsic Lyα luminosity function of galaxies (black solid line) together with the predicted apparent luminosity function for our bubble (red lines), web (blue lines), and web-bubble (green lines) models with different values of $\langle f_{\mathrm{HI}} \rangle_V$. Fig. 2.6 shows that

- The predicted luminosity function decreases with $\langle f_{\mathrm{HI}} \rangle_V$ as naturally expected, because more neutral hydrogen in the universe increases the overall opacity to Lyα photons.

- The relative abundance of large-scale bubbles and small-scale absorbers is a key factor to estimate the observed Lyα luminosity function. The bubble, web, and web-bubble models predict almost identical luminosity functions for vastly different $\langle f_{\mathrm{HI}} \rangle_V$. For example, a bubble model with $\langle f_{\mathrm{HI}} \rangle_V = 0.676$ (B2) gives rise to a luminosity function that is practically indistinguishable from that of a web model with $\langle f_{\mathrm{HI}} \rangle_V = 0.032$ (W3) or of a web-bubble model with $\langle f_{\mathrm{HI}} \rangle_V = 0.373$ (B1+W2). This was first pointed out by [44].

- The presence of small-scale absorbers inside ionized bubbles provides an opacity additional to that from the neutral patches between large-scale bubbles. This is clear

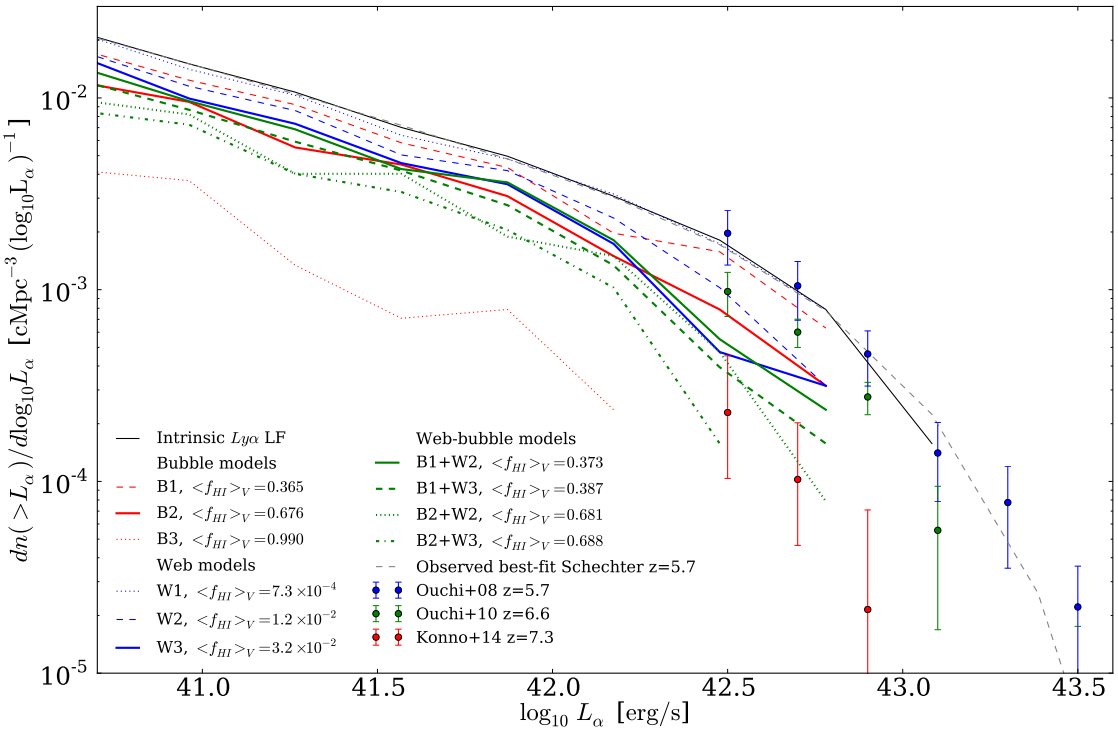

Figure 2.6: Intrinsic (black line) and observed differential Lyα luminosity functions at $z = 7$ as expected for bubble (red lines), web (blue lines), and web-bubble (green lines) reionization models. The observed data points of Ouchi et al. (2008, 2010) and Konno et al. (2014) are shown as blue, green, and red points. The best-fit Schechter function to the observations at $z = 5.7$ is shown as the gray dashed line. For each model, the line style refers to a different value of the parameters, as indicated in Table 2.1. The figure shows the degenerate impact of large-scale H I patches and small-scale absorbers on the Lyα LF.

comparing e.g. the LFs from B1 (dashed red line) to those from B1+W2 (dotted green) or B1+W3 (dashed green).

- Web models with $\langle f_{\rm HI}\rangle_V \sim 10^{-2}$ correspond to bubble models with $\langle f_{\rm HI}\rangle_V \sim 10^{-1}$. Table 2.1 indicates that this requires $\Gamma \leq 10^{-14}$ s^{-1}. For example, the red dashed (B1) and blue dashed (W2) lines in Fig. 2.6 show that $\Gamma_{-12}(z=7) = 0.01$ is needed for a web model to produce a LF similar to that of a bubble model with $\langle f_{\rm HI}\rangle_V \sim 0.4$. This is in agreement with [300].

- Comparing the simulations to the observations of [324, 325] and [234], we conclude that at $z = 6.6$, $40\% \lesssim \langle f_{\rm HI}\rangle_V \lesssim 70\%$ for the bubble model, $\langle f_{\rm HI}\rangle_V \sim 1\%$ for the web model, and $\langle f_{\rm HI}\rangle_V \lesssim 40\%$ for a web-bubble model. At $z = 7$ we have instead $70\% \lesssim \langle f_{\rm HI}\rangle_V \lesssim 99\%$, $\langle f_{\rm HI}\rangle_V \gtrsim 3\%$, and $40\% \lesssim \langle f_{\rm HI}\rangle_V \lesssim 70\%$, respectively. The inferred H I fraction thus highly depends on the reionization model adopted.

While the aim of this chapter is to present a proof of concept and we defer to future work more rigorous and precise constraints on the H I fraction, these results are in excellent agreement with existing work ([44, 300, 77]) and underline the importance of understanding the precise ionization structure of the IGM during the EoR in terms of both large-scale bubble features and small-scale absorbers. In the following, we use the simulations described in § 2.4 and the analytic formalism outlined in § 2.3 to gain more insight into the Lyα RT and the inference of $\langle f_{\rm HI}\rangle_V$ from observed Lyα-emitting galaxies.

2.5.2 The Red Damping Wing in Bubble Models

We now analyse the Lyα red damping wing opacity to quantify the impact of large-scale H I patches on the visibility of Lyα-emitting galaxies. The red lines in the top panel of Fig. 2.7 show the mean transmission $\exp(-\tau_{bub}^{\rm eff})$ as a function of Δv for three different values of $\langle f_{\rm HI}\rangle_V$ (B1, B2, and B3 from left to right). We evaluate the effective optical depth directly as an average of $e^{-\tau_{bub}}$ using line-of-sight skewers from galaxies extracted from the simulations. The shaded region indicates the 1σ dispersion $\sigma^2_{\tau_{bub}} = \langle (e^{-\tau_{bub}} - e^{-\tau_{bub}^{\rm eff}})^2 \rangle$. We have used 1185 lines-of-sight, i.e. equivalent to the number of galaxies in the simulation box.

The damping wing becomes more opaque with increasing neutral fraction and decreasing Δv. The opacity varies significantly between different lines-of-sight as indicated by the large dispersion of $\sigma_{\tau_{bub}} \sim 0.2$.

Comparison to the Analytic Model

To see how well the red damping wing can be captured by the analytic approximation, in Fig. 2.8 (left) we compare the results from our B2 model to those obtained using equation (2.12) with $\alpha_1 = 0.5$ and $R_* = 1.7,\ 3.0,\ 5.0,\ 10.0 h^{-1}$cMpc. Note that the case with $R_* = 1.7$ (thickest black dashed line) represents the Schechter function fit to the simulated $P(R_1)$ distribution. Fig. 2.8 (right) shows that the Schechter function fit is indeed a good

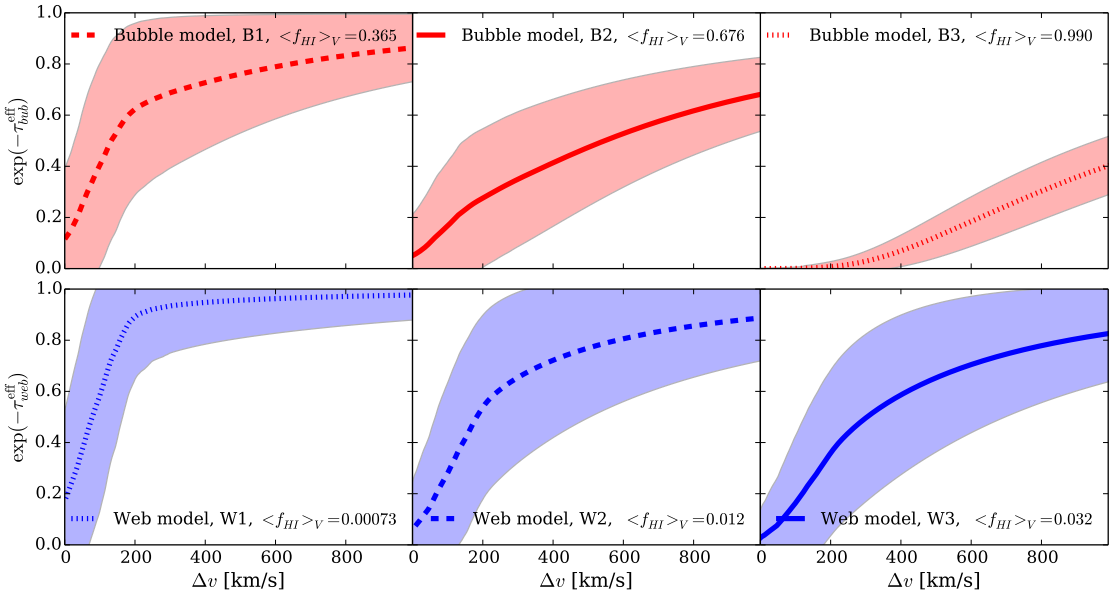

Figure 2.7: Effective optical depth (lines) extracted at $z = 7$ from simulations of the bubble (B1, B2 and B3; *top panels*) and web (W1, W2 and W3; *bottom panels*) models. The shaded regions are the corresponding 1σ dispersion.

approximation to the simulation, in which the distance to the near-side of the closest H I patch peaks at $\sim 5h^{-1}$cMpc from a galaxy.

The comparison in Fig. 2.8 (left) clearly indicates that the analytic model is too crude to capture the red damping wing behaviour found in the simulations, and systematically overestimates the optical depth, although the bubble size distribution is modelled reasonably well. The discrepancy highlights that the opacity is coming indeed from the neutral gas distributed among multiple H I patches, rather than in a single large H I patch, as assumed in the analytic model of equation (2.12). This in fact leads to an overestimate of the neutral gas and thus of the opacity.

In addition, the single large H I patch approximation is also responsible for a different shape of the damping wing, because the optical depth scales as Δv^{-1} (e.g. [307]). On the other hand, the presence of multiple ionized bubbles in the simulations makes the medium more transparent, and hence the damping wing profile steeper. This implies that, unless the analytic approximation is improved to take into account the complex ionized bubble distribution, (semi-)numerical simulations of patchy reionization are required to properly model the Lyα opacity in the diffuse neutral IGM[20].

[20]One obvious improvement of the analytic model would be to introduce an outer radius R_2 for the first diffuse neutral patch, and construct a PDF for R_2 which can then be included into equation (2.12) to give

$$e^{-\tau^{\rm eff}_{bub}(\nu_e)} \approx \int P(R_1)dR_1 \int e^{-\tau_{patch}(\nu_e,R_1,R_2)} P(R_2|R_1), \tag{2.22}$$

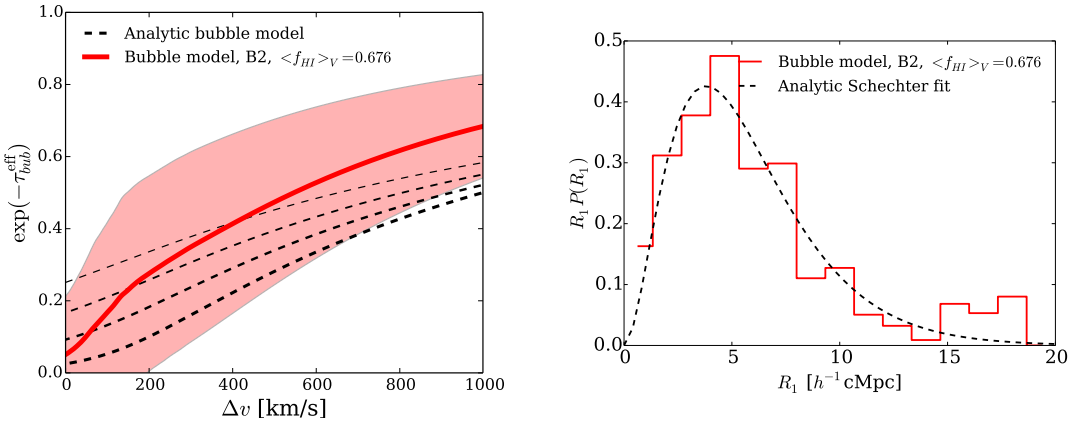

Figure 2.8: (Left) Comparison between the analytic and simulated bubble model. The black dashed lines are the effective optical depth from the analytic approximation (equation (2.12)) with a fixed slope $\alpha_1 = 0.5$ and $R_* = 1.7,\ 3.0,\ 5.0,\ 10.0h^{-1}$cMpc (lines from bottom to top). The discrepancy between the simulation and the analytic model is due to the large single H I patch approximation used in the latter. See text for detail. (Right) Distribution of the distance, R_1, to the near-side of H I patches from Lyα-emitting galaxies. The red line shows the result for the bubble model simulation B2. The black dashed line is the Schechter function fit to the simulated $P(R_1)$, where the best-fit parameters are $\alpha_1 = 0.49$ and $R_* = 1.7h^{-1}$cMpc.

2.5.3 The Red Damping Wing in Web Models

The ensemble of small-scale absorbers can also form a damping wing feature in the effective optical depth towards Lyα galaxies as shown in the bottom panel of Fig. 2.7, where the blue lines and the shaded areas refer to the mean transmission $\exp(-\tau^{\rm eff}_{web})$ and to the 1σ dispersion $\sigma^2_{\tau_{web}} = \langle (e^{-\tau_{web}} - e^{-\tau^{\rm eff}_{web}})^2 \rangle$ as a function of Δv.

Similarly to the bubble model, Fig. 2.7 indicates that the damping wing in web models becomes more opaque with increasing neutral fraction and decreasing Δv. Neutral fractions $\langle f_{\rm HI} \rangle_V \sim 10^{-2}$ (W2 and W3), i.e. much higher than the one in the post-reionized universe (which is $\langle f_{\rm HI} \rangle_V \sim 10^{-4}$), are required to produce a $\sim 60-80\%$ reduction of Lyα visibility at $\Delta v = 600$ km s^{-1}. On the other hand, the effective optical depth in W1 (which has a neutral fraction closer to $\sim 10^{-4}$) is $e^{-\tau_{\rm eff}} > 0.9$ at $\Delta v = 600$ km s^{-1}, i.e. it hardly affects the Lyα visibility[21]. The scatter around the effective optical depth is again large, with $\sigma_{\tau_{web}} \sim 0.2$.

where $P(R_2|R_1)$ denotes the conditional probability of R_2, given R_1. We have started to include such improvement in our model. However, due to the difficulty in finding an analytic fitting function for $P(R_2|R_1)$, we have deferred this to a future work.

[21] This justifies the calibration of the intrinsic model discussed in § 2.4.3. However, the left bottom panel of Fig. 2.7 shows that $\exp(-\tau^{\rm eff}_{web}) \sim 0.6$ at $\Delta v \sim 100$ km s^{-1} even with a neutral fraction as small as the one in the post-reionized universe. For galaxies that have $\Delta v < 200$ km s^{-1} the impact of small-scale absorbers at $z < 6$ should therefore be taken into account.

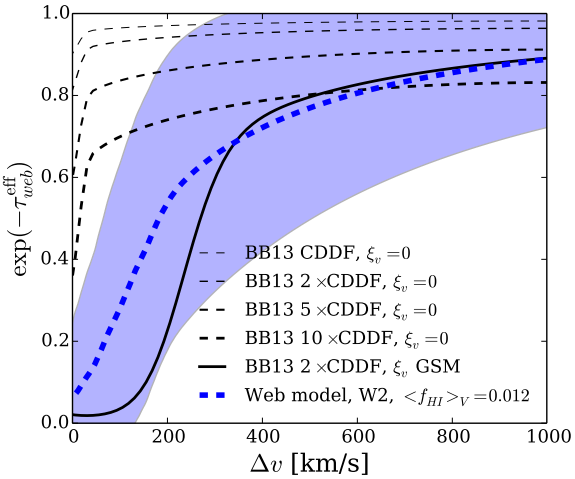

Figure 2.9: Comparison between the analytic and simulated web model. The black dashed lines are the effective optical depth from the analytic approximation (equation (2.14) using the BB13 CDDF without the velocity-space galaxy-absorber correlation function, i.e. $\xi_v = 0$. The black solid line instead includes a Gaussian streaming model (GSM) for ξ_v. This figure demonstrates the importance of the velocity-space galaxy-absorber correlation function.

Finally, a comparison between the effective optical depth in the web and bubble models (e.g. B2 vs. W3 in Fig. 2.7) shows that small-scale absorbers can produce a profile and scatter of the red damping wing similar to those of the bubble models. This explains the similarity in the Lyα LFs observed through the large-scale bubbles and small-scale absorbers.

Comparison to the Analytic Model

Fig. 2.9 compares the simulation and the analytic effective optical depth described by equation (2.14) in § 2.3.2. The black dashed lines refer to the analytic model without the effect of clustering and velocity field, i.e. $\xi_v = 0$ in equation (2.14), while the black solid line uses the Gaussian streaming model for ξ_v, i.e. equation (2.15). The analytic model employs a factor of 2-10 boost to the extrapolated CDDF fit by Becker & Bolton ([18]) at $z \simeq 7$ (hereafter BB13 CDDF) to mimic the rapidly increasing abundance of small-scale absorbers. Our fiducial value is 2 × CDDF (see § 2.5.3 a discussion on the reason of this choice).

Fig. 2.9 shows clearly that we cannot reproduce the results from our simulation by *only* changing the CDDF amplitude, while the agreement is much better if we simultaneously change the CDDF amplitude *and* include the galaxy-absorber correlation function in velocity-space (see § 2.5.3 for the reason of the discrepancy at $\Delta v < 200$ km/s). In other words, both the abundance of small-scale absorbers and their velocity-space clustering around galaxies play a key role in determining the Lyα visibility. In the following

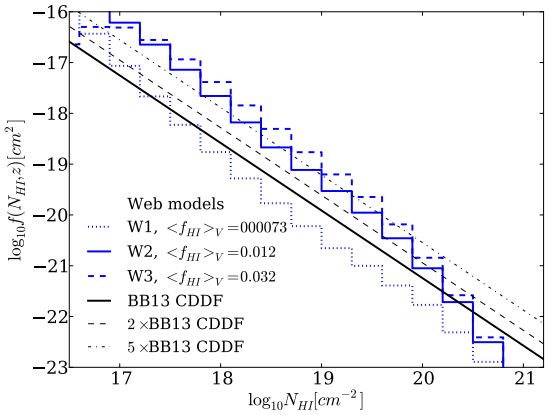

Figure 2.10: H I Column Density Distribution Functions (CDDFs) extracted from web model simulation W1 (blue dotted line), W2 (blue solid), and W3 (blue dashed). The black lines refer to the extrapolation to $z \simeq 7$ of the Becker & Bolton (2013) fit (BB13 CDDF; solid), multiplied by a factor of 2 (dashed) and 5 (dotted).

sections, we discuss in more detail the impact of (*i*) changing the CDDF (§ 2.5.3) and (*ii*) the galaxy-absorber clustering (§ 2.5.3).

CDDF and $N_{\rm HI}$-Dependence of the Optical Depth

We first justify the artificial boosting factor of the power-law CDDF adopted in the analytic model. Fig. 2.10 compares the CDDF obtained in our web model simulations[22] to the BB13 CDDF with a factor of 1, 2 and 5 boost. The adopted boosts broadly mimic the increase in simulated CDDF amplitude due to lower photoionization rate/higher neutral fraction ($\Gamma = 10^{-14}$ s^{-1} and 5×10^{-15} s^{-1} for W2 and W3), although the slope is not properly reproduced. The fiducial choice of 2 (corresponding to $\Gamma \sim 10^{-14}$ s^{-1}) approximately represents the CDDF amplitude in the range $10^{19}{\rm cm}^{-2} < N_{\rm HI} < 10^{20.7}{\rm cm}^{-2}$, which gives the highest contribution to the red damping opacity.

This predominance can be clearly seen in Fig. 2.11, which shows the ratio between the analytic effective optical depth from absorbers with column density below $N_{\rm HI}^{max}$ and below $N_{\rm HI}^{max} = 10^{21.3}{\rm cm}^{-2}$, in a case with[23] $\xi_v = 0$. More than 80 per cent of the optical depth redward of line centre ($\Delta v = 300$ km/s) comes from absorbers with $N_{\rm HI} > 10^{19}{\rm cm}^{-2}$, because of their prominent damping wings. On the other hand, at $\Delta v = -300$ km/s (i.e. blueward of the line resonance) lower column density absorbers with $N_{\rm HI} < 10^{18}{\rm cm}^{-2}$ can contribute ~ 50 per cent to $\tau_{web}^{\rm eff}$ via resonant absorption.

[22]We have computed the CDDF by taking the projected column density over 10 cells. The highest $N_{\rm HI}$ bins ($\log_{10} N_{\rm HI}/{\rm cm}^{-2} \sim 21.3$ and 20.6) are about ~ 0.3 dex larger than those calculated with a single cell, but the numbers converge for larger projected lengths. The effect is minor in the other bins.

[23]The inclusion of the velocity-space correlation function, for example inflowing low column density absorbers, would enhance the contribution of lower column density absorbers to the optical depth.

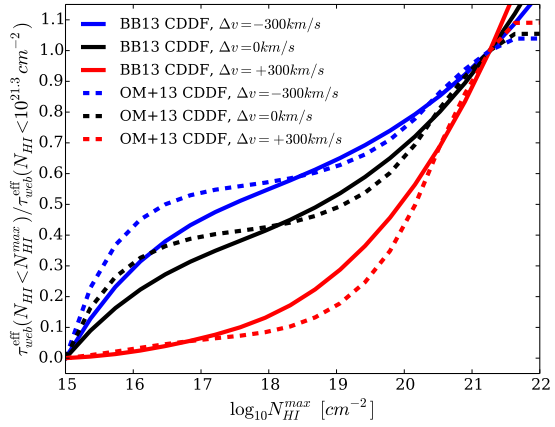

Figure 2.11: Cumulative contribution to $\tau^{\rm eff}_{\rm web}$ as a function of the maximum cut off column density $N^{max}_{\rm HI}$. Three different colours correspond to the optical depth redward (Δv = 300 km/s, red line), at line centre (Δv = 0 km/s, black line), and blueward (Δv = −300 km/s, blue line). The two functional forms of H I CDDF by Becker & Bolton (2013) (solid lines) and O'Meara et. al. (2013) (dashed lines) are plotted. This figure shows that the optical depth redward of line centre, i.e. red damping wing, is mostly sensitive to strong H I absorbers with $N_{\rm HI} > 10^{19}$ cm^{-2}, whereas the optical depth at smaller Δv is increasingly more affected by weaker H I absorbers with $N_{\rm HI} < 10^{17-19}$ cm^{-2}.

This strong dependence of the optical depth on the column density of absorbers is insensitive to the assumption about the shape of the CDDF, as shown by a comparison between the solid and dashed lines in Fig 2.11, which refer to models using a CDDF from Becker & Bolton ([18]) and O'Meara et al. ([320]), respectively. In both cases $N_{\rm HI} \gtrsim 10^{19}$cm^{-2} absorbers dominate the red damping wing, with a difference of only ~ 10 per cent.

Hence, the red damping wing opacity mainly depends on the abundance of strong H I absorbers, e.g. super-LLSs and DLAs, around Lyα-emitting galaxies. Their rapid increase (stronger than what expected from a simple extrapolation to $z > 6$ of lower-z CDDF) provides a large red damping wing opacity.

Galaxy-Absorber Correlation Function in Velocity-Space

The galaxy-absorber correlation function in velocity space, ξ_v, is another key factor in the formation of the red damping wing. In the Gaussian streaming model of equation (2.15), ξ_v depends both on (*i*) the real-space correlation function, $\xi(r_{12})$, and (*ii*) the galaxy-absorber pairwise mean velocity field $\langle v_{12}(r_{12})\rangle$, and pairwise velocity dispersion $\sigma_{12}(r_{12})$.

The simulated real-space galaxy-absorber correlation function at $z = 7$ is shown in Fig. 2.12 (left), together with the LBG-DLA correlation function observed by Cooke et al. ([90, 91]) at $z \sim 3$. The simulated $\xi(r_{12})$ is obtained by correlating the position of the galaxies and of the cells with $x_{\rm HI} > 0.9$ (which represent for us self-shielded absorbers) using

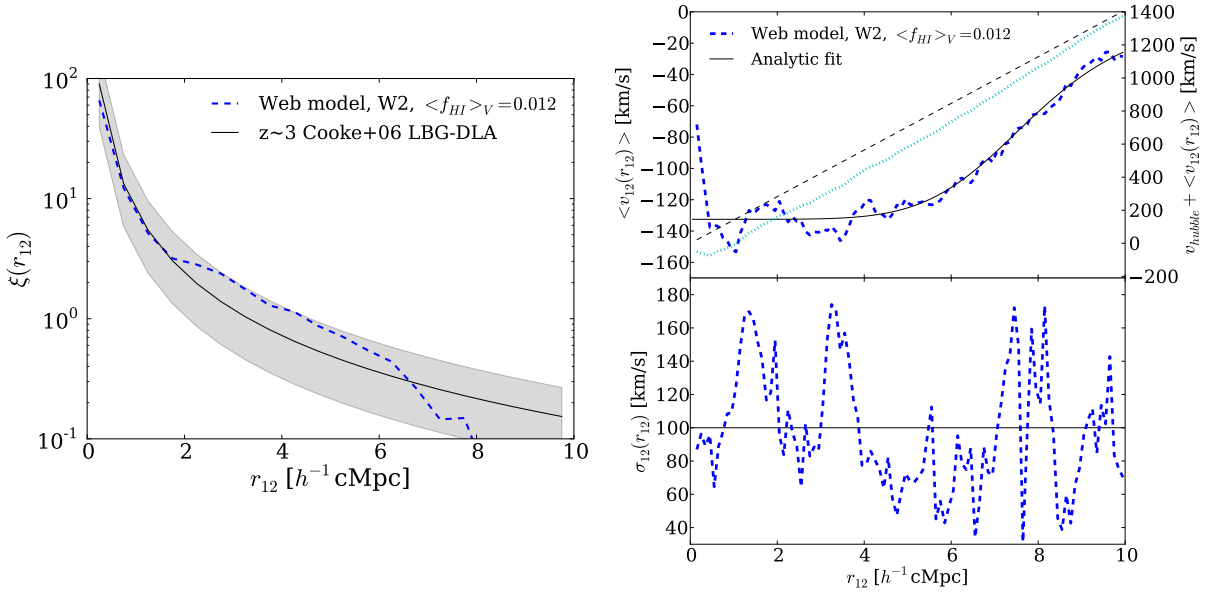

Figure 2.12: (Left) Simulated real-space galaxy-absorber correlation function at $z = 7$ (dashed blue line) and best-fit to the LBG-DLA cross-correlation function observed by Cooke et al. (2006b) at $z \sim 3$ (solid black). The latter is $\xi(r_{12}) = (r_{12}/r_0)^{-\gamma}$, where $r_0 = 3.32 \pm 1.25h^{-1}$cMpc, and $\gamma = 1.74 \pm 0.36$. The shaded region spans the upper and lower 1σ errors in the observed correlation length for a fixed slope $\gamma = 1.74$. The figure demonstrates the clustering of small-scale absorbers around galaxies. (Right) Galaxy-absorber mean pairwise velocity (*top panel*) and pairwise velocity dispersion (*bottom panel*) at redshift $z = 7$. The blue line (with left y-axis) shows the proper mean pairwise peculiar velocity $\langle v_{12}(r_{12})\rangle$ and pairwise velocity dispersion $\sigma_{12}(r_{12})$ between absorbers and galaxies in simulation W2, while the solid black lines are the fits adopted in the analytic calculation with the Gaussian streaming model. The cyan line corresponding to the right y-axis of the top panel refers to the total proper pairwise velocity, $v_{hubble} + \langle v_{12}(r_{12})\rangle$, and the black dashed line is the Hubble flow velocity $v_{hubble} = H(z_s)r/(1 + z_s)$. This figure shows the presence of cosmological inflow of absorbers onto galaxies, which slows down the total outflowing velocity including the Hubble flow.

the $\xi(r_{12}) = DD/RR - 1$ estimator ([104]). Clustering of self-shielding gas in the vicinity of Lyα-emitting galaxies is clearly important, and the simulated real-space correlation function appears (maybe surprisingly) similar to its lower-redshift observed counterpart. We thus adopt the Cooke et al. ([91]) correlation function for our Gaussian streaming model in Fig. 2.9.

The mean pairwise velocity between Lyα-emitting galaxies and absorbers defined above is shown in the top panel of Fig. 2.12 (right), both in terms of the proper peculiar velocity $\langle v_{12}(r_{12})\rangle$ (blue lines) and of the total proper velocity $H(z_s)r_{12}/(1+z_s) + \langle v_{12}(r_{12})\rangle$ (cyan lines). The solid black line is the best-fit curve to the mean pairwise velocity, $\langle v_{12}(r_{12})\rangle = -v_{in}/[1+(r_{12}/r_v)^{\gamma_v}]$ where $v_{in} = 133$ km/s, $r_v = 6.3h^{-1}$cMpc and $\gamma_v = 6.2$. This is adopted to evaluate the Gaussian streaming model in Fig. 2.9. For simplicity, rather than using a fit to the curve, we assume a constant pairwise velocity dispersion equal to its mean, i.e. $\sigma_{12} = 100$ km/s.

As shown in Fig. 2.9, the impact of the galaxy-absorber correlation function in velocity-space provides an additional boost of effective optical depth relative to the model with $\xi_v = 0$. In fact, the enhanced clustering of absorbers around galaxies (Fig. 2.12 (left)) renders the IGM more opaque. Furthermore, the cosmological inflow of absorbers onto galaxies (Fig. 2.12 (right)) causes a departure from the Hubble flow in the immediate surroundings of galaxies and enhances the velocity-space clustering (the slower the total outflow velocity in the proper unit is, the more opaque to Lyα photons the gas becomes, as it is less redshifted out of resonance). This can increase $\tau_{web}^{\rm eff}$, preferentially at the lower Δv. Thus, the effective optical depth including a velocity-space galaxy-absorber clustering is larger and steeper than the one including only a change in the CDDF amplitude (with $\xi_v = 0$).

Overall, Fig. 2.9 shows that the simulation and the analytic model agree at $\Delta v > 400$ km/s, while our analytic approximation overestimates the opacity at $\Delta v < 300$ km/s, probably because we assume that the same galaxy-absorber correlation function applies to the full column density range of absorbers. This may lead to low column density absorbers with a $\xi(r_{12})$ which is too large. To address this issue, it is necessary to investigate in more detail the column density dependent clustering, the pairwise velocity field with outflow, and/or the effect of photoionization from the central galaxy.

2.5.4 Lyα Red Damping Wing in Web-Bubble Models

The top panels of Fig. 2.13 show the effective optical depth in the hybrid web-bubble models directly calculated from the simulations, together with the 1σ dispersion of optical depth among different lines-of-sight. Not surprisingly, the red damping wing becomes more opaque towards higher neutral fractions, and the scatter from sight-line to sight-line is large. The red and blue lines show the contributions to the total simulated optical depth from the bubble and web models used to build each web-bubble model, while the black lines indicate the optical depth that we obtain by simply adding the two contributions, i.e. $\tau_{bub}^{\rm eff} + \tau_{web}^{\rm eff}$. A comparison between the optical depth extracted from the web-bubble simulations and a sum of the optical depths extracted from the corresponding bubble and

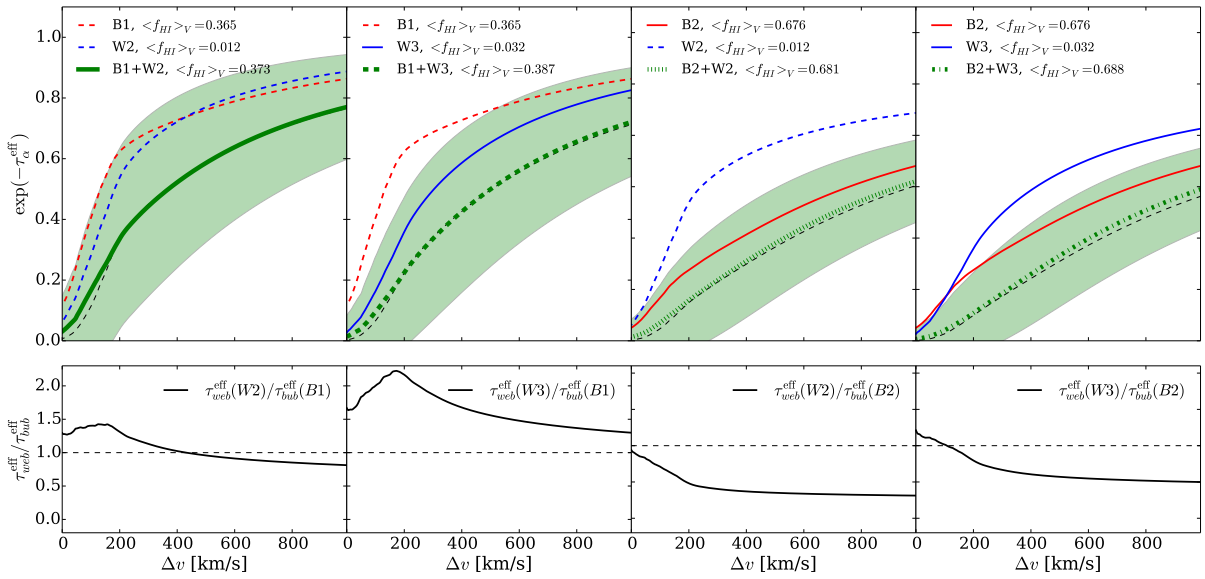

Figure 2.13: (*Top panels*) Lyα effective optical depth extracted from the hybrid web-bubble model simulations (green lines; B1+W2, B1+W3, B2+W2, and B2+W3 from left to right), together with the corresponding 1σ dispersion (shaded regions). The red and blue lines are the optical depth extracted from the bubble and web models used to construct the web-bubble models, and the black lines show the sum of these two contributions, i.e. $\tau_{bub}^{\rm eff}+\tau_{web}^{\rm eff}$. (*Bottom panels*) Ratio of effective optical depth between web and bubble models used for the corresponding web-bubble models. This shows the impact of large-scale bubbles and small-scale absorbers on the total optical depth as a function of Δv.

web models indicates that the two agree very well[24].

The bottom panels of Fig. 2.13 show the relative contribution of small-scale absorbers and large-scale H I patches to the total damping wing opacity. We find that:

- In all our web-bubble models, neither component dominates the total optical depth, as $\tau_{web}^{\rm eff}/\tau_{bub}^{\rm eff} \sim 0.5 - 1.5$.
- The relative contribution depends on the intrinsic Lyα line shift. The small-scale absorbers contribution increases with decreasing Δv because their opacity is enhanced by the inflow onto galaxies (see § 2.5.3). On the other hand, the H I patches are typically located at a distance $\sim 5-10h^{-1}$cMpc from Lyα-emitting galaxies (see Fig. 2.8 (right)), where the Hubble flow already dominates the total velocity. Therefore, they are more prominent at larger Δv.

The above two points underline the importance of correctly modelling small-scale absorbers within the large-scale bubble morphology. This section concludes our discussion on the average impact of large-scale neutral patches (§ 2.5.2) and small-scale absorbers (§ 2.5.3) on the Lyα red damping wing opacity in a unified web-bubble framework.

2.5.5 Probability Distribution Functions for $\mathcal{T}_{\rm IGM}$

Fig. 2.14 shows the distribution of the IGM transmission factor, $\mathcal{T}_{IGM}$, along the line-of-sight to a Lyα-emitting galaxy as a function of the host halo mass for models B2, W3, and B1+W2. These models have been chosen because they have a similar LF (see Fig. 2.6) and effective optical depth (see Figs. 2.7 and 2.13), and therefore a similar average Lyα visibility. The black lines are the average IGM transmission factor $\langle \mathcal{T}_{IGM}(M_h) \rangle$ for each halo mass bin.

In the bubble model plotted in the top panel of Fig. 2.14, $\langle \mathcal{T}_{\rm IGM}(M_h) \rangle$ increases with M_h, as massive [small] haloes typically reside in large [small] ionized bubbles (in the highest mass bins the trend is reversed because of the poor statistics). At the same time, there exists a population of lower mass haloes clustered around the more massive ones, which is therefore also embedded within large ionized bubbles. This explains the large scatter exhibit by $\mathcal{T}_{\rm IGM}$ for low halo masses. Furthermore, in bubble models sight-lines to most (if not all) galaxies pass through H I patches, meaning that the intrinsic luminosity of most galaxies is reduced, and explaining the *unimodality* of the $\mathcal{T}_{IGM}$ distribution (something that was pointed out previously by [213] and [300]).

As in web models self-shielding absorbers cluster around the more massive haloes (§ 2.5.3), $\langle \mathcal{T}_{\rm IGM}(M_h) \rangle$ decreases with increasing M_h. The still present large scatter in the distribution now appears to be *bimodal*, with a peak at $\mathcal{T}_{IGM} \sim 1$ and a second one at $\mathcal{T}_{IGM} \sim 0$. These peaks correspond to cases in which a line-of-sight intersects an absorber

[24] A slight discrepancy arises because the simple sum counts twice the neutral gas outside ionized bubbles (in the form of H I patches in bubble models and small-scale absorbers in web models), while in the simulations small-scale absorbers are present only within ionized bubbles by construction. The simple sum is thus expected to result in a slightly higher optical depth.

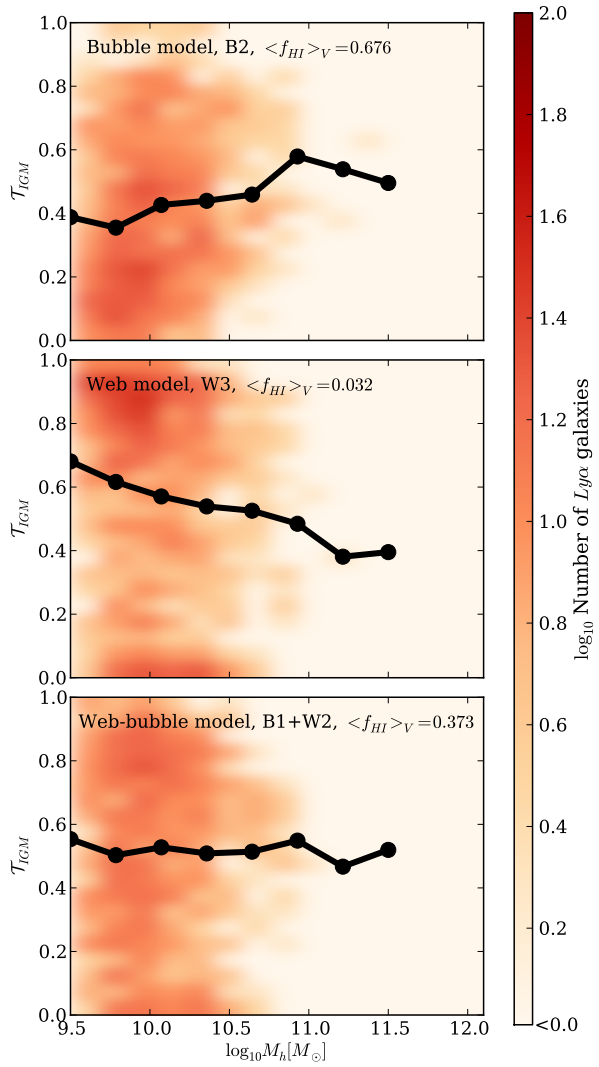

Figure 2.14: IGM transmission factor $\mathcal{T}_{IGM}$ along the line-of-sight to a Lyα-emitting galaxy as a function of the host halo mass M_h. The panels refer to the bubble model B2 (top), the web model W3 (middle), and the web-bubble model B1+W2 (bottom). The three models have a similar Lyα LF. The black lines are the average IGM transmission factors $\langle\mathcal{T}_{IGM}\rangle$ in each halo mass bin. The colour indicates the number of Lyα-emitting galaxies at each location of the $M_h - \mathcal{T}_{IGM}$ map, which is a 20(log) $\times$ 20(linear) grid in the range 10^9 M$_\odot$ $\leq M_h \leq 10^{12.5}$ M$_\odot$ and $0 \leq \mathcal{T}_{IGM} \leq 1$. The figure shows that the $\mathcal{T}_{IGM}$-PDF is unimodal for a bubble model and bimodal for a web model.

or not. Differently from what happens in the bubble model where the intrinsic luminosity of all galaxies is reduced, here a suppression is [is not] present depending on whether a small-scale absorber is [is not] aligned with a galaxy, hence the bimodality. Our results are consistent with those by Mesinger et al. ([300]), who also find that a bimodal distribution is a characteristic of the attenuation by small-scale absorbers.

In hybrid web-bubble models the IGM transmission factor is a product of large-scale bubbles and small-scale absorbers. Because of the different mass-dependence of $\mathcal{T}_{\rm IGM}$ in the two models, the total IGM transmission factor here depends only weakly on M_h, and no clear unimodality or bimodality in the distribution is visible. For example, the sight-lines present in the web model with $\mathcal{T}_{IGM} \sim 1$ are now more opaque due to the absorption from the H I patches between large-scale bubbles.

It is therefore clear that the conditional $\mathcal{T}_{IGM}$-PDF at a given halo mass, $P(\mathcal{T}_{IGM}|M_h)$, or in short the $\mathcal{T}_{IGM} - M_h$ relation, differs for web, bubble and web-bubble models. In the next section, we search for observational signatures of this variation in the intergalactic environment.

2.5.6 Simultaneously Constraining the H I Fraction and the Topology of Reionization

We now examine the prospect of observationally constrain the global H I fraction and the topology of reionization simultaneously by combining various statistics of Lyα emitting galaxies.

The Equivalent Width Distribution

Fig. 2.15 shows the cumulative probability distribution of the rest-frame equivalent width (REW), following the method of [121],

$$P(> \mathrm{REW}) = \int_0^1 P_{\rm intr}(> \mathrm{REW}/\mathcal{T}_{IGM}) P(\mathcal{T}_{IGM}) d\mathcal{T}_{IGM}, \tag{2.23}$$

where the intrinsic REW distribution is $P_{\rm intr}(> \mathrm{REW_{intr}}) = \exp(-\mathrm{REW_{intr}}/\mathrm{REW_c})$, with $\mathrm{REW_c} = 50$ Å ([121]) and $\mathrm{REW_{intr}} = \mathrm{REW}/\mathcal{T}_{IGM}$. The probability distribution function of the IGM transmission factor, $P(\mathcal{T}_{IGM}) \propto \int P(\mathcal{T}_{IGM}|M_h) \frac{dn(>M_h)}{dM_h} dM_h$, is constructed from the simulations.

In all models, the observed REW distribution is decreased in comparison to the intrinsic one by an amount which increases with the H I fraction. Similarly to what observed for the Lyα luminosity function, a degeneracy is present between web and bubble models, with, for example, B1 and W2 providing similar REW distributions.

However, the degeneracy can be partially broken if the REW distribution is combined with the Lyα LF. In fact, while models B2, W3 and B1+W2 are degenerate in Lyα LF (see Fig. 2.6) they produce distinguishable observed REW distributions. Although this is not always the case (for example, B1 and W2 show similar curves both in the Lyα LF and

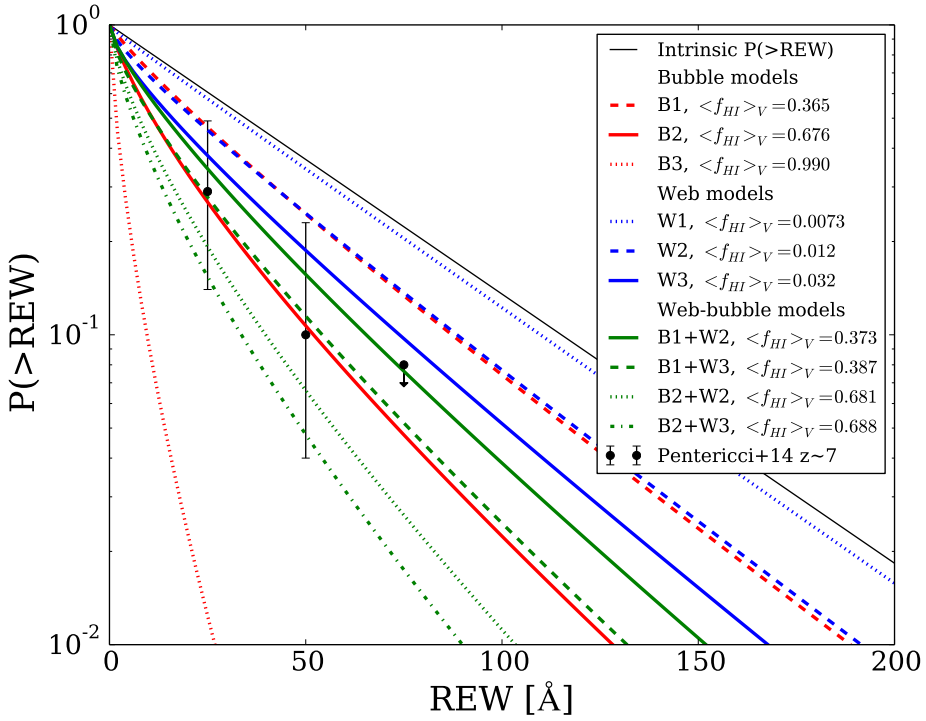

Figure 2.15: Cumulative probability distribution of the rest-frame equivalent width at $z = 7$. The black line is the intrinsic REW distribution and the coloured lines refer to the observed REW distributions predicted from simulations: bubble model B1 (red dashed), B2 (red solid) and B3 (red dotted); web model W1 (blue dotted), W2 (blue dashed) and W3 (blue solid); web-bubble model B1+W2 (green solid), B1+W3 (green dashed), B2+W2 (green dotted) and B2+W3 (dotted-dashed). The black circles are the observation of Pentericci et al. (2014) without interloper correction (if the interloper correlation is taken into account the data points can be higher by $\sim 20\%$).

the REW distribution), such a combined analysis offers a test to differentiate reionization models.

The argument above can be better understood by noting that the observed Lyα LF and REW distribution depend differently on the $\mathcal{T}_{IGM} - M_h$ relation. To see this, we first express the Lyα LF in terms of $P(\mathcal{T}_{IGM}|M_h)$ as

$$\frac{dn(> L_\alpha^{obs})}{dL_\alpha} = \int P(L_\alpha^{obs}|M_h)\frac{dn(> M_h)}{dM_h}dM_h, \tag{2.24}$$

where

$$P(L_\alpha^{obs}|M_h) = \int_0^1 P_{\rm intr}(L_\alpha^{obs}/\mathcal{T}_{IGM}|M_h)P(\mathcal{T}_{IGM}|M_h)d\mathcal{T}_{IGM}. \tag{2.25}$$

$P_{\rm intr}(L_\alpha|M_h)$ is the intrinsic conditional probability distribution of the Lyα luminosity given a halo mass[25]. A comparison between equations (2.23) and (2.25) shows a different dependence on $P(\mathcal{T}_{IGM}|M_h)$[26]. This is because the Lyα LF is constructed from Lyα selected LAEs, while the REW-PDF is constructed from continuum selected galaxies. In fact, [123] and [181] have shown that selection by Lyα line flux enhances the contribution of UV-faint galaxies (at fixed Lyα flux), which are absent from continuum selected samples. As such UV-faint galaxies should preferentially reside in low mass haloes, this difference in selection function would introduce a different dependence in the $\mathcal{T}_{IGM} - M_h$ relation that may lead to a drop in the observed Lyα LF different from the one in the REW distribution.

Hence, a combined analysis of Lyα LF and REW distribution may allow to constrain the H I fraction and the topology of reionization. We can already do this analysis. The upper limit at REW = 75 Å slightly favours the bubble or web-bubble models with $\langle f_{\rm HI}\rangle_V \sim 68\%$ or $\sim 37\%$. If we include this constraint, the neutral fraction is favored to be of order of tens of per cent. This constraint is very weak because of a large uncertainty due to the interloper contamination. Moreover, the same observations favour bimodal quenching of the Lyα visibility, which is associated with web-models. This argument simply illustrates that a combined analysis of Ly LF and REW-PDF can shed light on the history and topology of reionization.

Lyα Fraction of Lyman-Break Galaxies

The power of such joint analysis can be strengthened once the M_{UV}-dependent Lyα fraction of LBGs measurement is included as well. The Lyα fraction of LBGs (hereafter $\mathcal{X}_{\rm Ly\alpha}$) is defined as the fraction of LBGs with a UV magnitude M_{UV} and Lyα REW greater than a given value. We generalize the method of Dijkstra et al. ([121], see also [123]) to calculate

[25]Explicitly, we use $P_{\rm intr}(L_\alpha|M_h) = \delta_D[L_\alpha - L_\alpha(M_h)]$ as we assume a one-to-one mapping between L_α and M_h based on the abundance matching technique.

[26]Note that equation (2.23) implicitly assumes that the intrinsic REW distribution is independent of halo mass. We can, of course, generalize this modelling to include the halo mass dependence, but because this in general differs from the one of the Lyα luminosity, the dependence of the two statistics on the $\mathcal{T}_{IGM} - M_h$ relation is expected to differ as well.

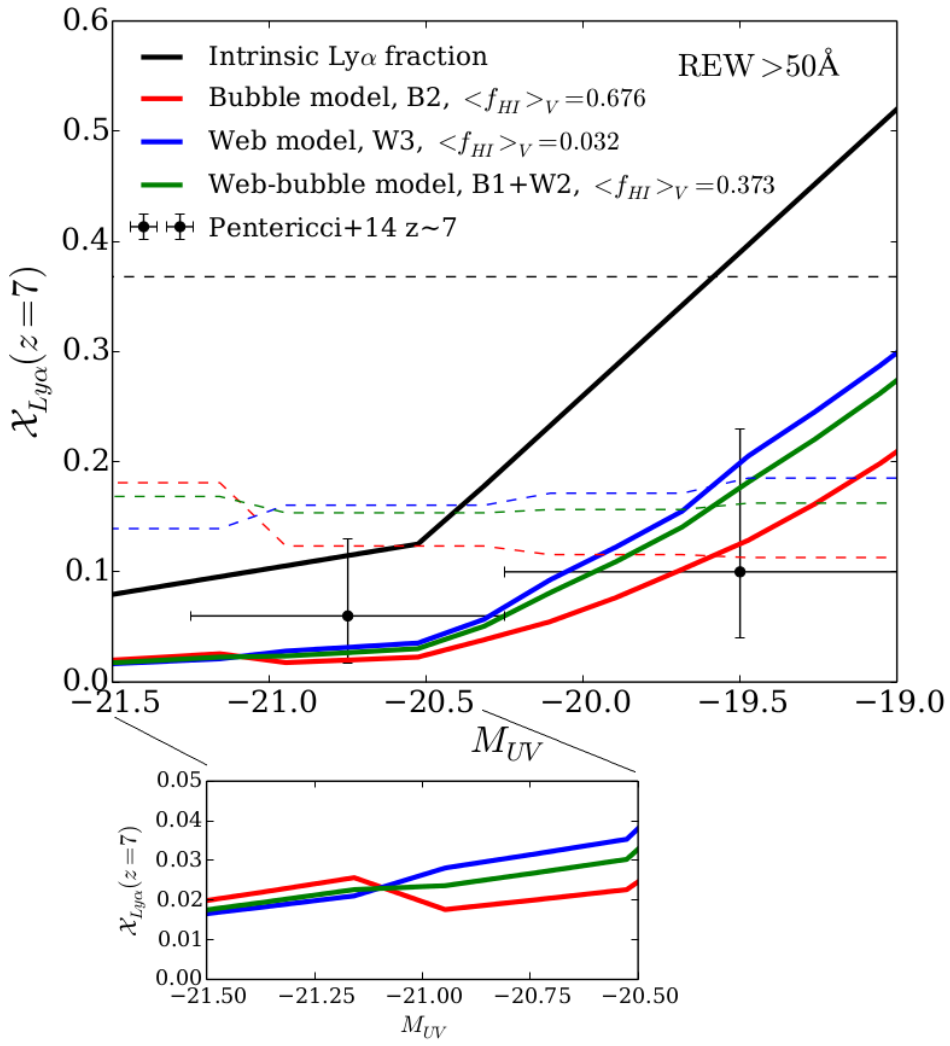

Figure 2.16: Lyα fraction of LBGs having REW $>$ 50 Å at $z = 7$ as a function of the UV magnitude M_{UV} of galaxies. The line style indicates the intrinsic fraction obtained with the M_{UV}-dependent (solid lines) and the uncorrelated (dashed lines) model (see text for more details), and the line colours refer to the intrinsic Lyα fraction (black), and to the observed Lyα fraction for models B2 (red), W3 (blue) and B1+W2 (green). The three models have a similar Lyα LF (Fig. 2.6). The black points are the observations of Pentericci et al. (2014), where the horizontal error bars indicate the bin size used. The smaller plot is a zoom-in of the Lyα fraction - M_{UV} relation to emphasize the upturn at UV-bright LBGs caused in model B2 by their larger Lyα visibility.

the Lyα fraction of LBGs as

$$\mathcal{X}_{\rm Ly\alpha}(>{\rm REW}|M_{UV}) = \int_0^1 \mathcal{X}^{\rm intr}_{\rm Ly\alpha}(>{\rm REW}/\mathcal{T}_{IGM}|M_{UV})P(\mathcal{T}_{IGM}|M_{UV})d\mathcal{T}_{IGM}, \quad (2.26)$$

where $\mathcal{X}^{\rm intr}_{\rm Ly\alpha}(>{\rm REW_{intr}}|M_{UV})$=$e^{-{\rm REW_{intr}}/{\rm REW_c}(M_{UV})}$ is the intrinsic Lyα fraction, ${\rm REW_c}(M_{UV})$ is a characteristic REW (see Appendix A.5 for more details), and the conditional $\mathcal{T}_{IGM}$ probability distribution function at a given M_{UV} is

$$P(\mathcal{T}_{IGM}|M_{UV}) = \int P(\mathcal{T}_{IGM}|M_h)P(M_h|M_{UV})dM_h. \quad (2.27)$$

We construct our intrinsic model assuming that UV-bright LBGs populate more massive haloes, and consider a case with a correlation between REW and M_{UV} (M_{UV}-dependent model) and one with no correlation (uncorrelated model). The M_{UV}-dependent model is our fiducial case because observations suggest that such correlation exists ([425, 215], but see [316]). More details are provided in Appendix A.5.

Fig. 2.16 shows the intrinsic and observed Lyα fractions at $z = 7$ for models B2, W3, and B1+W2, which all produce a similar Lyα LF. Two main features emerge:

- the bubble model shows an *upturn* of Lyα fraction at UV-bright LBGs (typically defined as those with $M_{UV} < -20.25$), while the web model shows a monotonic decrease of Lyα fraction for increasing UV-bright LBGs[27]. This qualitative change in the shape of the M_{UV}-dependent Lyα fraction is robust against different intrinsic models of REW.

- In the M_{UV}-dependent model (solid lines), the drop in the observed Lyα fraction compared to the intrinsic one is larger for UV-faint LBGs ($M_{UV} > -20.25$) than for UV-bright LBGs in *all* models. The common expectation that the drop of Lyα fraction of UV-faint LBGs is larger than the one of UV-bright LBGs occurs *only* in bubble models ([321]) is true *only if* the intrinsic REW and M_{UV} are uncorrelated (dashed lines).

The upturn of the M_{UV}-dependent Lyα fraction can be understood as an imprint of the $\mathcal{T}_{IGM} - M_h$ relation (see § 2.5.5). In fact, because UV-bright LBGs in bubble models are more likely to be surrounded by large ionized bubbles, the probability that their Lyα emission is visible (i.e. that they are associated to larger $\mathcal{T}_{IGM}$) is higher than for UV-faint LBGs. On the other hand, in web models the small-scale absorbers cluster more strongly around UV-bright LBGs, lowering their Lyα visibility. The upturn of Lyα fraction, therefore, does not happen in web models. As a consequence, the qualitative change in the *shape* of the M_{UV}-dependent Lyα fraction can be used as an indicator of the (possible) presence of large-scale bubbles.

[27]Note that the downturn in the bubble model B2 at $M_{UV} \lesssim -21.15$ is due to the poor statistics, similarly to what observed in the top panel of Fig. 2.14.

On the other hand, a drop of the Lyα fraction for UV-faint LBGs larger than for UV-bright LBGs cannot be used as a decisive evidence of patchy reionization. In fact, while in the uncorrelated case (dashed lines) we indeed see a larger drop for UV-faint LBGs only for the bubble model, in the M_{UV}-dependent case (solid lines) such drop is visible for all models. The simplest explanation for this is that because, to first order approximation, the neutral IGM suppresses the Lyα emission by re-scaling the characteristic REW as $\langle \mathcal{T}_{IGM} \rangle \mathrm{REW_c}(M_{UV})$ (see also Appendix A.5), UV-faint galaxies (with an intrinsically larger $\mathrm{REW_c}$) experience a larger reduction in number above a given REW than the UV-bright galaxies (with intrinsically small $\mathrm{REW_c}$) do.

In summary, the analysis of the M_{UV}-dependent Lyα fraction of LBGs provides a powerful diagnostic tool to characterize the impact of large-scale bubbles and small-scale absorbers when properly interpreted. Hence, when combined with the Lyα LF, it offers an opportunity to constrain the H I fraction and the topology of reionization simultaneously. While the aim of the present chapter is to highlight the potential of this diagnostics, we plan to use it more extensively in a future study.

2.6 Discussion & Conclusions

The visibility of Lyα-emitting galaxies during the Epoch of Reionization is controlled by both diffuse H I patches in the IGM, and small-scale self-shielding absorbers around galaxies. It is therefore important to correctly include small-scale absorbers inside large-scale ionized bubbles. In this work we have explored the impact of both large-scale bubbles and small-scale absorbers on the visibility of the population of Lyα-emitting galaxies at $z > 6$, using a powerful combination of an analytic approach and hydrodynamical simulations, which covers the full range of models explored in recent investigations ([214, 44, 300, 77]). We have considered the IGM Lyα RT in three different classes of IGM ionization structure, namely (*i*) the bubble model, where only large-scale ionized bubbles due to patchy reionization are present, (*ii*) the web model, where only small-scale absorbers are considered, and (*iii*) the web-bubble model, which includes both small-scale absorbers and large-scale bubbles.

Our main conclusions are:

- The observed Lyα LF evolution from $z = 5.7$ to $z \sim 7$ requires a neutral fraction $\langle f_{\rm HI} \rangle_V \sim 60 - 80\%$ in bubble models, $\langle f_{\rm HI} \rangle_V \gtrsim 1 - 3\%$ in web models, and $\langle f_{\rm HI} \rangle_V \sim 30 - 70\%$ in web-bubble models.

- A sole analysis of the Lyα luminosity function or of the distribution of rest frame equivalent width cannot put a stringent constraint on the reionization history. The Lyα LF function and the REW-PDF can be equally suppressed in bubble, web, and web-bubble models, yet with very different global H I fractions. Hence, there is a fundamental degeneracy between the ionization structure of the IGM and the global H I fraction inferred from Lyα surveys (see § 2.5.1).

- We showed in § 2.5.6 that a joint analysis of the Lyα LF and the REW-PDF of LBGs can improve the constraints on the neutral fraction by breaking the degeneracy with the topology of reionization.

- The Lyα fraction of LBGs can be a powerful diagnostic to study the relative importance of large-scale H I patches and small-scale absorbers in the IGM. We caution that a drop in Lyα fraction that is larger for UV-faint LBGs than for UV-bright LBGs (as in [321]) can be reproduced with web and web-bubble models, and does not provide exclusive evidence for patchy reionization. Instead, we argue that the shape of the M_{UV}-dependent Lyα fraction may provide more insight into the topology of reionization (see e.g. Fig. 2.16).

  For example, an upturn of Lyα fraction for UV-bright LBGs can be caused by large-scale ionized bubbles, but also by an increase in the UV background around UV-bright galaxies, which reduces the abundance of small-scale absorbers. Interestingly, this upturn may already have been observed at $4.5 < z < 6$ ([425]), and may reflect large fluctuations in the UV background. These fluctuations have been proposed to explain observations of the cumulative effective optical depth distribution at $z \gtrsim 5$ in the spectra of high-redshift QSOs ([22, 73]).

- Our analytic formalism shows that the Lyα damping wing opacity from small-scale absorbers is highly influenced by the clustering and the pairwise velocity field of galaxy-absorber pairs (see § 2.5.3). Absorbers with $N_{\rm HI} > 10^{19}$ cm^{-2}, i.e. super-LLS/DLAs, provide the largest contribution to the the red damping wing at $\Delta v >$ 300 km s^{-1}, while lower column density absorbers are important at smaller Δv. Understanding the galaxy-absorber correlation functions and their velocity fields can improve the robustness with which the reionization history can be constrained using Lyα emitting galaxies. Direct observational constraints on H I CDDF and galaxy-absorbers correlation function (and as a function of $N_{\rm HI}$) can therefore be very useful. A possible approach is to extend to the range $3 < z < 7$ the survey strategy that searches for Lyα-emitting galaxies in the foregrounds of high-redshift QSOs, similar to the observation of [91], Keck Baryonic Structure Survey ([378, 456]), and VLT LBG Redshift Survey ([92]). This observational strategy is already within reach at $z \sim 5.7$ ([114]).

- We showed that the total effective optical depth in web-bubble models can be written as the sum of those in web and bubble models, i.e. $\tau_\alpha^{\rm eff} \approx \tau_{bub}^{\rm eff} + \tau_{web}^{\rm eff}$ (see § 2.5.4). This is an important result as fast semi-numeric simulations can be used to generate $\tau_{bub}^{\rm eff}$. These simulations can then be complemented with (improved) analytic or possibly empirical prescriptions for $\tau_{web}^{\rm eff}$ (as in § 2.5.3) to efficiently generate more realistic web-bubble models.

- Web, bubble and web-bubble models produce different $\mathcal{T}_{\rm IGM}$-PDFs (§ 2.5.5). Bubble

models show a unimodal $\mathcal{T}_{\rm IGM}$-PDF, while small-scale self-shielding absorbers in the web-model have a bimodal $\mathcal{T}_{\rm IGM}$-PDF. The modality of the hybrid web-bubble model depends on which component dominates the IGM opacity. Pentericci et al. ([343]) have provided observational evidence for bimodal quenching of Lyα flux (see [453, 452] for details on the procedure). Our results imply that bimodal quenching indicates an influence of small-scale absorbers on the Lyα visibility (also see [300]), which is opposite to the common interpretation.

In conclusion, in this chapter we have shown that a joint analysis of different statistics of Lyα emitting galaxies (e.g. Lyα LF, REW distribution, Lyα fraction of LBGs, correlation function), can break degeneracies associated with individual probes. It should therefore be possible to constrain simultaneously the global H I fraction *and* the reionization topology, when armed with a suit of models of reionization in which both large-scale bubble morphology and small-scale absorbers are included.

Chapter 3

Reionization by Galaxies and QSOs and the Thermal State of the IGM

In Chapter 3 I present a detailed physics of reionization driven by galaxies and QSOs. Understanding the physical states of the IGM is important to exploit the observations of QSO absorption spectra as well as future 21cm tomography.

3.1 Backgrounds

In the current paradigm of extragalactic astronomy, the universe has undergone two major reionization epochs. First, the reionization of H I to H II (H I reionization) and He I to He II (He I reionization) are thought to occur around $z \sim 6-20$. The second reionization epoch of helium, He II $\rightarrow$ He III (He II reionization), occurred around $z \sim 2-4$. These two epochs correspond to the time when first galaxies and the early descendants have assembled and the peak of cosmic star formation and quasar activities. Therefore, star forming galaxies and QSOs play an important role in characterizing the physical state of the intergalactic medium (IGM) over a wide range of cosmic history $z \sim 2-20$. In turn, this means that astrophysics of the IGM is critical to understand the full reionization process, and galaxy formation and their interaction with the environment.

The ionization state of the IGM provides a primary information for understanding the role of galaxies and QSOs during the reionization epoch. Current observations of the H I Lyα Gunn-Peterson trough suggest that hydrogen reionization is largely completed by $z \sim 6$ (e.g. [140]). However, the nature of the sources that drive the epoch of hydrogen reionization is still unknown. A concordance picture is to assume (undetected) faint star forming galaxies as the main reionizing agents with an escape fraction of ionizing photons as large as ~ 20 per cent (e.g. [373]). At lower redshift, observations of the He II Lyα Gunn-Peterson trough suggest that He II reionization is mostly completed by $z \sim 2.5$ ([484, 433]). Unlike H I reionization, He II reionization can be driven by already detected population of QSOs ([170]). Simulations calibrated with the observed QSO luminosity function can grossly reproduce the observed properties of He II Lyα forest ([286, 86, 87]).

Hydrogen and helium reionization also impacts on the thermal states and their large-scale fluctuations in the IGM ([244, 283, 178]). This means that the reionization epoch can be constrained by measuring the IGM temperature. The temperature measurements from Lyα forests ([283, 253]) suggest that the end of H I reionization should be lower than $z \simeq 9$ ([442, 41, 363]), as well as He II reionization is extended over $2 \leq z \leq 4.8$ ([19]). Furthermore, the IGM temperature in the QSO near zone at $z \sim 6$ suggests that the onset of He II reionization may be as early as the formation of first QSOs at $z > 6$ ([40]).

Although many authors have studied the impact of QSOs on the ionization and thermal states of the IGM (e.g. [262, 70, 486, 447, 254, 103]), the role of QSOs during H I reionization at $z > 6$ still remains unclear. While it has been argued that QSOs alone cannot be responsible for hydrogen reionization based on constraints from the unresolved X-ray background ([118, 382, 383, 185]) and the decreasing number density of high-z QSOs ([290]), recent re-investigations suggest an possible important contribution from QSOs (e.g. [171, 170, 260]).

Furthermore, theoretical predictions of how QSOs influence the ionization and thermal states of the local environment have not always agreed. For example, there seems to be no consensus about whether a QSO imprints a distinctive ionization and thermal structures on the IGM by producing a very large ionized region (e.g. [143, 499, 444]), whether it is indistinguishable from the contribution of galaxies ([7, 254, 103]), whether the difference in the spectra of galaxies and QSOs produces a different shape of ionization front (I-front) ([498, 236]), or whether the sphericality of an ionized region may serve as discriminator of how the region is created ([103]). These theoretical discrepancies must be resolved to interpret QSO absorption spectra and upcoming 21cm observations. In other words, the correct theoretical understanding is vital to place constraints on the role of galaxies and QSOs in driving reionization.

Understanding the impact of reionization driven by galaxies and/or QSOs on the physical state of the IGM boils down to an accurate modeling of the effects of (i) the spectral shape of QSOs and galaxies (i.e. using a multi-frequency radiative transfer (RT)), (ii) galaxies surrounding a QSO (cosmological N-body/hydrodynamical simulations), (iii) anisotropic propagation of I-front (3D simulation), and (iv) coherent treatment of both UV and X-ray photons as well as secondary ionization (UV/X-ray physics). Previous works focused on the above aspects separately; for example, 1D RT simulations by Thomas & Zaroubi ([444]) and Kramer & Haiman ([236]) focused on (i) and (iv), whereas the 3D RT simulation by Datta et al. ([103]) which included both galaxies and QSOs, but without modelling the thermal structure, focused on addressing (i), (ii) and (iii)[1]. At the moment,there is no radiative transfer calculation reported, addressing all the four effects together in a single simulation.

In this chapter we therefore investigate the ionization states of both hydrogen and helium and the thermal state of the IGM in a single high-z QSO environment by performing a suite of multi-dimensional, multi-frequency radiative transfer simulations post-processing

[1]Feng et al. ([143]) and Keating et al ([221]) have also conducted RT simulation post-processing cosmological hydrodynamical simulation, but without simultaneously accounting both galaxies and QSO.

a cosmological hydrodynamical simulation (thus, satisfying (i)-(iv)). This is the most detailed calculation of this kind to date, at least to our knowledge. This work distinguishes itself from previous works as we present the case-by-case analysis of hydrogen and helium reionization as well as the thermal state in the QSO environment including surrounding galaxies, X-ray and secondary ionization. This large suite of RT simulations aims at providing insights into the underlying physical mechanisms responsible for controlling the physical state of the IGM around QSOs, and how all processes collectively shape the state of the IGM. Our work is an important first step for full cosmological reionization simulations including both galaxies and multiple QSOs for future work.

This chapter is organized as follows. First we describe the simulation setup in § 3.2. § 3.3 presents the results of hydrogen and helium reionization. § 3.4 discusses the thermal states in the QSO environments. We compare our results with previous works in § 3.5. Observational implications are briefly discussed in § 3.6. Conclusions are then presented in § 3.7.

3.2 Simulations

In the following we will describe the simulations employed in this paper. Hereafter we assume the cosmological parameters: $\Omega_\Lambda = 0.74$, $\Omega_m = 0.26$, $\Omega_b = 0.0463$, $h = 0.72$, $n_s = 0.95$ and $\sigma_8 = 0.85$.

3.2.1 Hydrodynamical simulations of the IGM

We have performed hydrodynamical simulations of the IGM with the smoothed particle hydrodynamics (SPH) code GADGET-3, which is an updated version of the publicly available code GADGET-2 ([420]). The size of the simulation box is $(50h^{-1}$ cMpc$)^3$, with 2×512^3 dark matter and gas particles, corresponding to a mass of $5.53 \times 10^7 h^{-1}\mathrm{M}_\odot$ and $1.20 \times 10^7 h^{-1}\mathrm{M}_\odot$, respectively. Haloes are identified using the friend-of-friends algorithm with a linking length of 0.2. The simulations are then centred on the most massive halo (with mass $1.34 \times 10^{10} h^{-1}\mathrm{M}_\odot$) at $z = 10$, which is modelled to host a QSO[2] (see § 3.2.4). All haloes are considered to harbor one galaxy per halo with the empirically motivated prescription described in § 3.2.3, which results in 7534 galaxies at $z = 10$.

The gas density and temperature are mapped onto Eulerian grids with 256^3 cells; in total 11 snapshots are present in the redshift range $z = 15 - 10$, with a $\Delta z = 0.5$ interval. These are used as input for the RT code `CRASH` (see § 3.2.2).

[2]We note that this is small compared with a typical mass of haloes inferred to host QSO. This is a necessary shortcoming to have higher resolution of the IGM and RT simulation (but see Keating et al [221] for a recent zoom-in simulation). To bypass this issue, we model the luminosity of a QSO based on $z \sim 7$ observation, instead of scaling the luminosity with halo mass.

3.2.2 Multi-frequency radiative transfer simulations

The outputs of the hydro simulations described above are post-processed with the Monte Carlo code CRASH ([81, 270, 269, 177]; Graziani, in prep), which, in its latest version, solves the cosmological radiative transfer of UV and X-ray photons[3] in a gas constituted by hydrogen and helium[4]. We refer the reader to the original papers for more details on the code. Here we assume a H and He number fraction of $X = 0.92$ and $Y = 0.08$.

The code self-consistently evolves the ionization states of hydrogen and helium and recalculates the gas temperature by solving for non-equilibrium chemistry and radiative heating and cooling processes. The photoionization heating of H I, He I, and He II, the recombination cooling, and the adiabatic cooling by cosmological expansion are self-consistently included within the RT calculation (see [270] for a more comprehensive list of heating/cooling mechanisms implemented in the code).

For our multi-frequency RT calculations, the source spectrum from 13.6 eV to 2 keV is sampled by 42 frequency bins using 29 bins for the UV range, with finer sampling near the ionization threshold of H I (13.6 eV), He I (24.6 eV) and He II (54.4 eV), and 13 bins to cover the soft X-ray (200 eV $< h\nu <$ 2 keV) range. The Monte-Carlo simulations cast 10^6 photon packets per galaxy and 2×10^8 per QSO. We have verified this by performing the convergence test.

The radiative transfer of X-ray photons and the effect of fast photo-electrons produced by high energy photons are described in detail in Graziani et al. (2016, in prep.). Secondary ionization effects have been modelled by many authors (e.g. [406, 100, 460, 160]). Although the results are qualitatively similar in the previous studies, there are some quantitative differences on the IGM thermal and ionization states. Here we adopt the model by [460] as it is the most recently reported model implemented in CRASH, and we refer the reader to Graziani et al. (2016, in prep.) for a more comprehensive comparative analysis.

A suite of RT simulations: overview

We create a suite of RT simulations to study the reionization models by galaxies and QSOs. The simulation suite contains 14 runs in total for our main analysis. We describe the overall procedure adopted for producing the RT simulation suite.

Our multi-frequency RT simulations consist of two steps: (1) Firstly, we perform three runs only with galaxy-type sources from $z = 15$ until $z = 10$ (see Table 3.1). (2) Secondly, starting from the final snapshot ($z = 10$) of the first step, we further evolve the simulations including only galaxies ('Galaxies only' runs) and including both galaxies and a QSO ('Galaxies+QSO' runs) (see Table 3.2). In addition, the second step also contains simulations including only a QSO ('QSO only' runs), starting from the fully neutral initial condition. These three sets of runs are designed to model the galaxy-driven, QSO-driven, and galaxy+QSO-driven reionization.

[3] We refer to photons in the energy range 13.6 eV-200 eV (200 eV-2 keV) as UV (soft X-ray) photons.
[4] While the code can handle the effect of metals, we do not include them here.

Table 3.1: Parameters used for the simulations of reionization history from galaxies in the range $z = 15 - 10$. From left to right, the columns refer to the simulation ID, UV photon parameter $f_{\rm UV}$, galaxy spectrum slope α_Q, and the start and end redshifts of RT calculation.

run ID	$f_{\rm UV}$	α_G	z
0.5R	0.05	3	$15 - 10$
R	0.1	3	$15 - 10$
2R	0.2	3	$15 - 10$

Table 3.2: Parameters used for the simulations at $z = 10$ described in the text. From left to right, the columns refer to the simulation ID, UV photon parameter $f_{\rm UV}$, galaxy/QSO spectrum slope α_G/α_Q, inclusion/exclusion of the reionization history (history) from galaxies in the range $z = 15 - 10$, a QSO, the X-ray photons and secondary ionization. The one highlighted in boldface (GAL_R+QSO_UVXsec) is our reference run.

run ID	$f_{\rm UV}$	α_G, α_Q	history	QSO	X-ray	secondary
Galaxies only						
GAL	0.1	3, n/a	off	off	off	off
GAL_0.5R	0.05	3, n/a	on	off	off	off
GAL_R	0.1	3, n/a	on	off	off	off
GAL_2R	0.2	3, n/a	on	off	off	off
QSO only						
QSO_UV	n/a	n/a, 1.5	off	on	off	off
QSO_UVX	n/a	n/a, 1.5	off	on	on	off
QSO_UVXsec	n/a	n/a, 1.5	off	on	on	on
QSO_obsc_UVXsec	n/a	n/a, 1.5(obscured)	off	on	on	on
Galaxies+QSO						
GAL_R+QSO_UV	0.1	3, 1.5	on	on	off	off
GAL_R+QSO_UVX	0.1	3, 1.5	on	on	on	off
GAL_0.5R+QSO_UVXsec	0.05	3, 1.5	on	on	on	on
GAL_R+QSO_UVXsec	**0.1**	**3, 1.5**	**on**	**on**	**on**	**on**
GAL_2R+QSO_UVXsec	0.2	3, 1.5	on	on	on	on
GAL_R+QSO_obsc_UVXsec	0.1	3, 1.5(obscured)	on	on	on	on

The three runs in the first step in Table 3.1 are performed with the three different total ionizing photon emissivities of galaxies. Our fiducial run is 'R' run (standing for Reference). The other two runs ('0.5R' and '2R') differ only in the total ionizing photons by a factor of 1/2 and 2. The source models of galaxies are described in § 3.2.3. The final snapshots are dumped at $z = 10$ as initial conditions for the restart simulation in the next step.

The second step contains 14 main runs to be analysed in § 3.3 and 3.4. All runs (re)start at $z = 10$ and are evolved for 10^8 yr. We use the output snapshot at 10^7 yr for our main analysis unless otherwise stated. In all three categories of the source models (Galaxies only, QSO only, and Galaxies+QSO models), the run IDs with '0.5R', 'R', and '2R' start from the final snapshots of the first step shown in Table 3.1 (the column 'history' in Table 3.2). For 10 runs with a QSO ('QSO only' and 'Galaxies+QSO' runs), a QSO is turned on at $z = 10$. They are performed with various combinations of source models and RT physics to study the impact of a QSO in a exquisite detail (see § 3.2.4 for QSO source model). For example, the runs with and without the previous reionization history from galaxies in a QSO environment allow us to study the effects of a QSO on the IGM in isolation and with surrounding galaxies. In addition to the variations of the source models, the main runs also contain the various combinations of RT physics (the columns 'X-ray' and 'secondary'). Runs with X-ray 'on' include the contribution from soft X-ray photons up to 2 keV, while X-ray 'off' artificially truncates the input spectra of sources at 100 eV. 'Secondary' indicates whether we switch on the secondary ionization effect. Our reference full run is GAL_R+QSO_UVXsec as indicated in boldface in Table 3.2. Our suite of RT simulations are summarized in Tables 3.1 and 3.2.

3.2.3 Source model: galaxies

To assign an ionizing photon luminosity to a galaxy we follow the method described in [80], and we refer the reader to this paper for more details. Briefly, the comoving ionizing photon emissivity of the entire galaxy population at $z > 6$ is modelled based on a pre-assumed global star formation rate density and the observations of the ionizing background from Lyα forest ([43]). The total ionizing photons are then distributed among all galaxies in a simulation box.

The comoving ionizing photon emissivity is given by (in units of photons s^{-1} cMpc^{-3}),

$$\dot{n}_{ion}(z) = 10^{50.89}\chi(z)\frac{\alpha_b + 3}{2\alpha}f_{\rm UV}, \tag{3.1}$$

where α and α_b are the extreme-ultraviolet power-law spectral index of the sources and of the ionizing background, which are assumed to be identical and equal to 3. The redshift dependence of star formation rate density is parameterized as $\chi(z) = \frac{\xi e^{\zeta(z-9)}}{\xi-\zeta+\zeta e^{\xi(z-9)}}$ with $\xi = 14/15$ and $2/3$. The normalization of the ionizing background is parameterized by a free parameter, $f_{\rm UV}$. This is introduced to allow flexibility in the model to explore the uncertainty of the ionizing background in a QSO environment. The fiducial value is $f_{\rm UV} = 0.1$, and we consider a variation of a factor of 2 around it, i.e. $f_{\rm UV} = 0.2, 0.05$. Figure 3.1

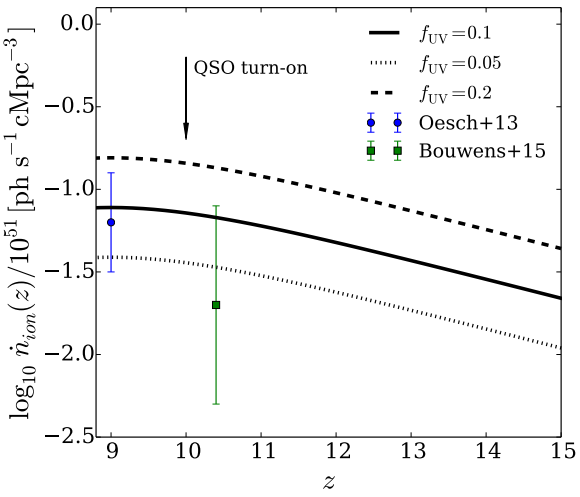

Figure 3.1: Total ionizing photon emissivity from galaxies as a function of redshift. The three model values of $f_{\rm UV} = 0.05$ 0.1, 0.2 (dotted, solid, and dashed curves) are compared with the values given by Becker et al. (2015) which are based on the observations of Lyman-break galaxies by Oesch et al. (2013) and Bouwens et al. (2015). The figure shows that our models are a reasonable choice to bracket the possible parameter space.

shows that the comparison of our ionizing photon emissivity with the values inferred from observations of Lyman-break galaxies ([20] based on [319] and [53]). They agree reasonably well. Given that the total ionizing photon emissivity in a QSO environment is very poorly constrained, the level of agreement shown in Figure 3.1 justifies our choice of $f_{\rm UV}$ parameter. All the run IDs with 'GAL_' in Table 3.2 and the runs in Table 3.1 adopt the three values of $f_{\rm UV} = 0.05, 0.1, 0.2$ to quantify the different contribution of galaxies to reionization.

To distribute the total ionizing photons among galaxies, we assume that each halo hosts one galaxy. The ionizing photon luminosity of a galaxy is assumed to scale linearly with the host halo mass, $\dot{N}_{ion}^{\rm GAL}(M_h) \propto M_h$. To recover the total ionizing photon emissivity when we add up all photons from galaxies, the ionizing photon luminosity, $\dot{N}_{ion}^{\rm GAL}(M_{h,i})$, of the i-th galaxy with halo mass, $M_{h,i}$, is given by (in units of photons s^{-1}),

$$\dot{N}_{ion}^{\rm GAL}(M_{h,i}) = \dot{n}_{ion}(z) V_{box} \frac{M_{h,i}}{\sum_{j=1}^{N_s} M_{h,j}}, \tag{3.2}$$

where V_{box} is the comoving volume of the simulation box and N_s is the total number of galaxies (haloes) in the simulation. $\dot{n}_{ion}(z)V_{box}$ is the total number of ionizing photons present at redshift z.

The spectrum of a galaxy is modelled with a power-law spectral energy distribution (SED), $L_\nu^{\rm GAL}(\nu) \propto \nu^{-\alpha_G}$, with spectral index $\alpha_G = 3$. The ionizing photon luminosity and SED is related as $\dot{N}_{ion}^{\rm GAL} = \int_{\nu_L}^{\infty} \frac{L_\nu^{\rm GAL}(\nu)}{h\nu} d\nu$ where $\nu_L = 3.29 \times 10^{15}$ Hz is the frequency at the Lyman limit. The spectrum is plotted in Figure 3.2 (solid blue line), together with a

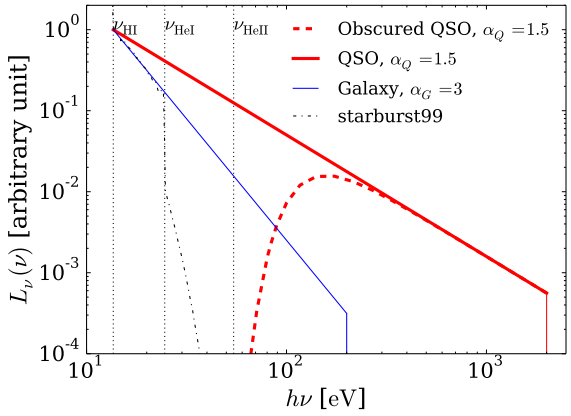

Figure 3.2: The model spectra of a QSO (red solid), obscured QSO (red dashed), and galaxy (blue solid). For a comparison, the black dash-dotted line shows a spectrum at 10 Myr after a burst of star formation with metallicity $Z = 0.001$ from starburst99. The vertical lines show the ionization thresholds of H I, He I, and He II from left to right.

spectrum at 10 Myr after a starburst with metallicity $Z = 0.001$ from starburst99 (black dotted line, [250]). While the H I-ionizing photon luminosity between $h\nu_{\rm HI}$ and $h\nu_{\rm HeI}$ is well approximated by the power-law, it drops abruptly at $> h\nu_{\rm HeI}$ due to the absorption by stellar atmosphere in the spectral synthesis model. The adoption of a simple power-law allows a more straightforward comparison with a QSO model (§ 3.2.4). Although our main conclusions are not affected by the power-law assumption, we are aware of the limitation. For example, the ionization of helium and heating by galaxies would be smaller if we used the population synthesized spectra.

3.2.4 Source model: QSO

For a QSO at $z = 10$, we assume an ionizing photon luminosity of

$$\dot{N}_{ion}^{\rm QSO} = 1.36 \times 10^{56} \text{ photons s}^{-1}. \tag{3.3}$$

As a reference, ULAS J1120+0641 QSO at $z = 7.085$ is estimated to have a central supermassive black-hole (SMBH) of mass $M_{\rm BH} = 2 \times 10^9$ M$_\odot$ and the ionizing photon luminosity of $\dot{N}_{ion}^{\rm QSO} = 1.3 \times 10^{57}$ photons s^{-1} ([313, 47]). We model the ionizing photon luminosity of $z = 10$ QSO by re-scaling the properties of ULAS J1120+0641. In our model QSO, we assume a smaller SMBH mass of $M_{\rm BH} = 2.1 \times 10^8$ M$_\odot$ because we are modelling a state of a higher redshift QSO environment at $z = 10$. Assuming the Eddington limit accretion rate, the bolometric QSO luminosity scales linearly with the SMBH mass, which translates to the linear scaling between an ionizing photon luminosity and a SMBH mass, $\dot{N}_{ion}^{\rm QSO} \propto M_{\rm BH}$. Therefore, the assumed SMBH mass of $z = 10$ QSO corresponds to the value in equation (3.3).

The QSO spectrum is assumed to have a power-law index $\alpha_Q = 1.5$, (red solid line in Figure 3.2),

$$L_\nu^{\rm QSO}(\nu) = L_0 \left(\frac{\nu}{\nu_L}\right)^{-\alpha_Q}, \tag{3.4}$$

where the normalization factor L_0 is determined from $\dot{N}_{ion}^{\rm QSO} = \int_{\nu_L}^{\infty}(L_\nu^{\rm QSO}/h\nu)d\nu$. This power-law index is identical to the one of ULAS J1120+0641 ([313, 47]).

In addition, we also create an obscured QSO model (red dashed line in Figure 3.2). Our model spectrum is obscured by a H I column density of $N_{\rm HI} = 10^{20}$ cm^{-2}. This H I column density is chosen as a reasonable fiducial value based on lower redshift observations of AGNs ([241, 458]) and the one assumed in a similar theoretical work ([236]). The ionizing photons are absorbed by a factor of $\exp[-\sigma_{\rm HI}(\nu)N_{\rm HI}]$ where $\sigma_{\rm HI}(\nu)$ is the H I-photoionization cross section ([128]). Therefore, the ionizing photon luminosity escaping to the IGM is given by

$$L_\nu^{\rm QSO,obsc}(\nu) = L_0 \left(\frac{\nu}{\nu_L}\right)^{-\alpha_Q} e^{-\sigma_{\rm HI}(\nu)N_{\rm HI}}. \tag{3.5}$$

The normalization factor L_0 and the power law index α_Q are the same as the (unobscured) QSO model. Note that in an obscured QSO model the ionizing photon luminosity is reduced to $\dot{N}_{ion}^{\rm QSO,obsc} = 4.33 \times 10^{54}$ photons s^{-1} because a large fraction of ionizing photons with $\nu < \nu_{\rm HeII}$ are absorbed before escaping into the IGM.

The lifetime of the QSO is assumed as $t_Q = 1 \times 10^7$ yr unless otherwise stated. This is chosen from the observationally allowed values 10^6 yr $< t_Q < 10^8$ yr at lower redshift ([268]). A different choice of luminosity and lifetime, of course, affects the simulation result. Because the recombination timescale is much shorter than the QSO lifetime, increasing lifetime and/or luminosity trivially increases the size of ionized regions as $R_{\rm I} \propto (\dot{N}_{ion}^{\rm QSO} t_Q)^{1/3}$. The temperature is expected to be insensitive to luminosity and lifetime, although the radius to which gas is photo-heated is increased with increasing luminosity and lifetime.

All the QSO source models shown in Table 3.2, except for the two cases marked as 'obscured', adopt the luminosity and spectrum of the unobscured QSO model.

3.3 Results

In this section we discuss the results of our suite of radiative transfer simulations for reionization of both hydrogen and helium driven by galaxies, a QSO, and galaxies+QSO.

Hydrogen and helium reionization: overview

The 3D visualization in Figure 3.3 summarizes the results for our full reference simulation (GAL_R+QSO_UVXsec). The figure shows clearly the very rich morphology of H II and He II/He III regions in the high-z QSO environment. First, the QSO and surrounding galaxies together form a central large H II region on scales of tens of mega-parsecs. The clustering of galaxies around a QSO is also responsible for producing many smaller H II

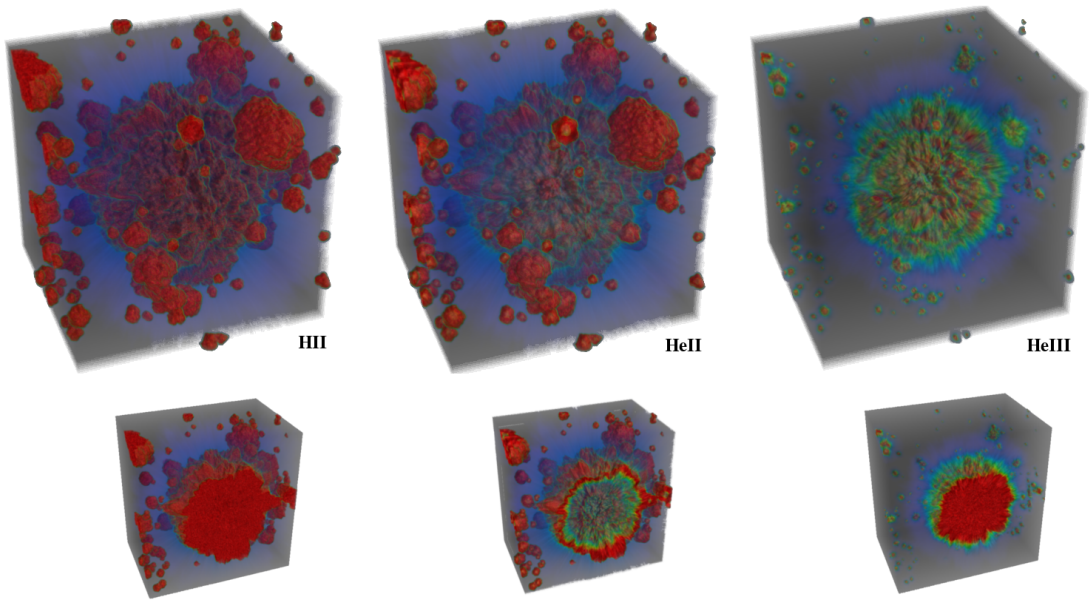

Figure 3.3: 3D visualization of the H II (left), He II (middle), and He III (right) fractions in the QSO environment, as produced by both the central QSO and the surrounding galaxies with full UV/X-ray physics at $z = 10$ (GAL_R+QSO_UVXsec). *Top panels*: the whole simulation box of $50h^{-1}$cMpc. *Bottom panels*: dissection through the central QSO.

regions. The X-ray radiation from the QSO partially ionizes the IGM beyond the sharp H II I-front. Next, the He II regions show a morphology similar to that of H II regions. The He II I-front around the central H II region is slightly broader than H II I-front because of the slightly larger mean free path of photons near He I ionization threshold from the QSO. The interior of the central QSO-host He II region is, however, drastically different from that of the H II region. The He II-ionizing photons ($h\nu > 54.4$ eV) from a QSO drive the second reionization of helium from He II to He III. This forms a cavity inside the He II region by further He III reionization. Finally, the morphology of He III regions is clearly different from those of H II and He II regions. It is dominated by a central large He III region, with only a few tiny He III regions around. This is because the ionization threshold of He II is high, the sources of He II reionization are dominated by QSOs. The He III I-front is even broader due to the larger mean free path of photons near He II ionization threshold. Overall, the reionization of hydrogen and helium around a QSO is driven by the concerted photoionization by UV and X-ray radiations from both the central QSO and the surrounding galaxies.

In the following (§ 3.3.1-3.3.4), we examine the details of hydrogen and helium reionization by varying the source models of galaxies and QSOs and the physics of radiative transfer.

3.3.1 H II and He II/He III regions in a galaxy- or QSO-driven reionization model

We first examine the simplest (but unrealistic) case in which either galaxies or a QSO drive reionization starting with a fully neutral initial condition. These two different source models impact differently the properties of H II and He II/He III regions.

Hydrogen

Figure 3.4 (top panels) shows slices of H II fraction, $x_{\rm H\,II}$, in the two models, GAL and QSO_UVXsec, where only galaxies or a QSO ionize a fully neutral IGM for a duration of 10^7 yr starting at $z = 10$. The difference in the morphology of H II regions between the two source models is clear: while a single large fully ionized region appears around the QSO, the numerous fainter galaxies produce many smaller H II regions. This morphological difference is caused by the fact that in the QSO model all the photon budget is concentrated in one single source rather than being distributed among many galaxies. In our models the QSO emits about 15× more ionizing photons than the galaxies.

In addition, the partially ionized region around a QSO is another prominent difference in the ionization structure. It is produced by the X-ray photons from a QSO. The detail of the impact of X-ray is discussed in § 3.3.2.

For a more quantitative comparison, in the top panels of Figure 3.5 we plot the distribution of the I-front radii[5], $R_{\rm I}^{\rm HII}$, produced by the QSO (red solid line) and the galaxies (blue dashed line) along 100 random lines-of-sight (LOSs) from the central source. While a central QSO alone can produce an H II region as large as $R_{\rm I}^{\rm HII} \approx 13h^{-1}$cMpc, galaxies give only a small I-front radius of $R_{\rm I}^{\rm HII} < 3h^{-1}$cMpc. Furthermore, a QSO-type source shows $\sim 5h^{-1}$cMpc dispersion of the H II I-front radii in different LOSs. Such dispersion of the I-front positions is caused by the density fluctuations of the IGM; while dense gas clumps cast shadows behind, I-fronts propagate unimpeded along underdense directions ([207, 221]).

Helium

The middle and bottom panels of Figure 3.4 show maps of He II and He III fractions. As for the H II fraction, a clear different morphology is visible between the galaxies only and QSO only models.

When reionization is driven by galaxies, the morphology of the He II and H II regions is qualitatively very similar. The reason for this is that, although our galaxy-type spectrum emits a factor of $L_\nu^{\rm GAL}(\nu_{\rm HeI})/L_\nu^{\rm GAL}(\nu_{\rm HI}) = (\nu_{\rm HI}/\nu_{\rm HeI})^3 \simeq 0.17$ less He I- than H I-ionizing photons, because the abundance of helium is lower by a factor of $Y/X \simeq 0.09$, the galaxies can drive He I reionization together with H I reionization. As a result, He I and H I have a similar reionization history and morphology. On the other hand, the He III regions are more

[5]The I-front radius is defined as the distance from the central source at which $x_{\rm H\,II}$ drops below 50% for the first time.

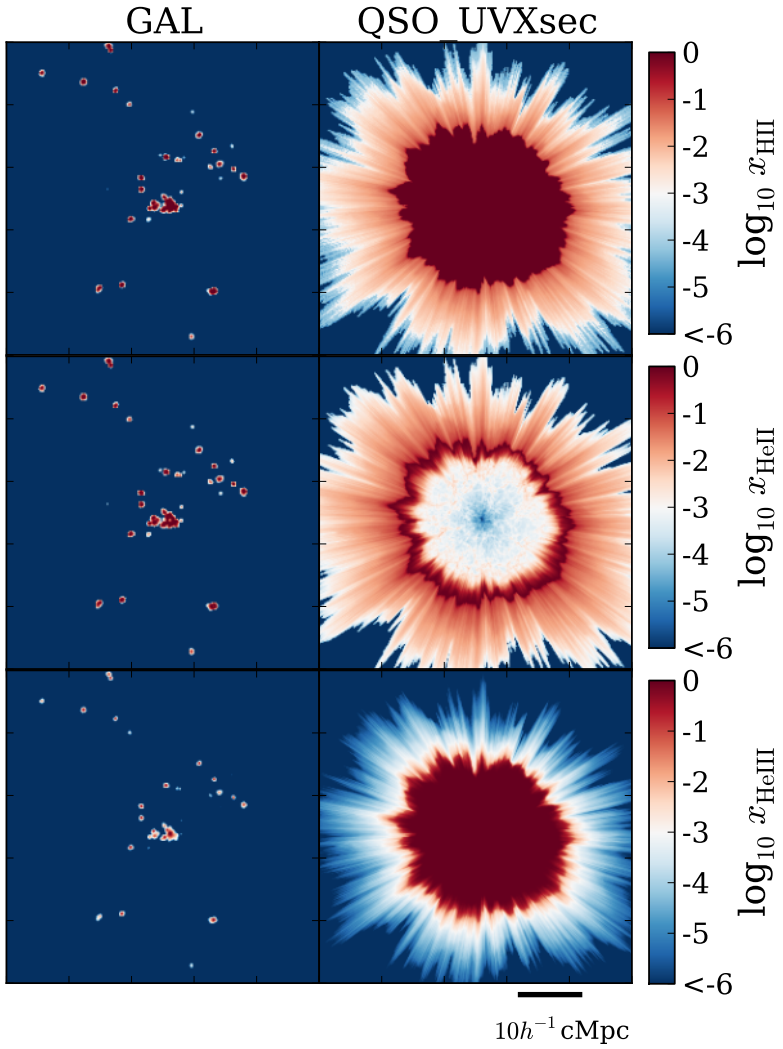

Figure 3.4: Maps of H II (top panels), He II (middle), and He III (bottom) fractions at $z = 10$ for model GAL (left column) and QSO_UVXsec (right). In both models the sources are turned on in a fully neutral medium at $z = 10$ and shine for 10^7 yr. The maps have a side length of $50h^{-1}$cMpc and the width of the slice is $195h^{-1}$ckpc.

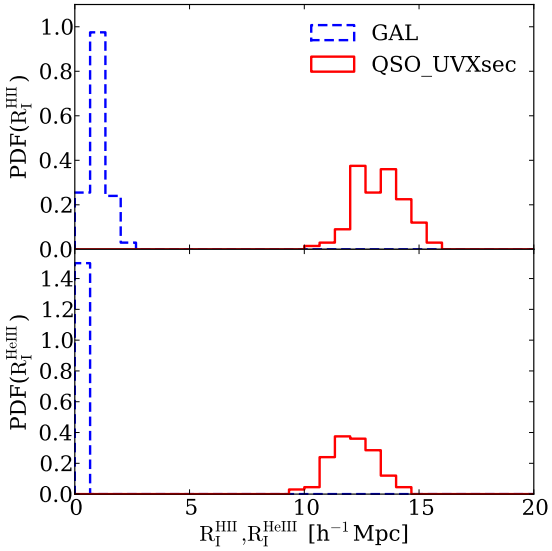

Figure 3.5: Distribution of the H II (top panel) and He III (bottom panel) I-front radii for models GAL (blue dashed) and QSO_UVXsec (red solid). We exclude the corresponding histrogram for He II I-fronts because the location is nearly identical to H II I-fronts.

compact. This is a consequence of the higher ionization threshold of He II, which results in a factor of $L_\nu^{\rm GAL}(\nu_{\rm HeII})/L_\nu^{\rm GAL}(\nu_{\rm HI}) = (\nu_{\rm HI}/\nu_{\rm HeII})^3 \simeq 0.016$ less He II- than H I-ionizing photons.

A QSO, on the other hand, produces a very distinctive He II region compared to its H I region, while the QSO He III region has a morphology similar to its H II region. A QSO-type spectrum produces a large amount of He I- and He II-ionizing photons per H I-ionizing photons; $L_\nu^{\rm QSO}(\nu_{\rm HeI})/L_\nu^{\rm QSO}(\nu_{\rm HI}) = (\nu_{\rm HI}/\nu_{\rm HeI})^{1.5} \simeq 0.4$ and $L_\nu^{\rm QSO}(\nu_{\rm HeII})/L_\nu^{\rm QSO}(\nu_{\rm HI}) = (\nu_{\rm HI}/\nu_{\rm HeII})^{1.5} \simeq 0.13$, respectively. As a result, a QSO drives He II reionization simultaneously to H I reionization. Because the He I reionization has not been completed by $z = 10$, the QSO He III I-front drives He II reionization immediately after He I reionization. This creates a cavity in the He II region, which is characteristic to the high-z QSO environment.

The distribution of the He III I-front radii is shown in the bottom panel of Figure 3.5. QSO He III I-fronts (red solid line) extends to $R_{\rm I}^{\rm HeIII} \approx 12h^{-1}$cMpc, which are lagging slightly behind the QSO H II I-fronts. On the other hand, galaxies (blue dashed line) has a negligibly small radius of He III I-front. The dispersion of the QSO He III I-front positions is as large ($\sim 5h^{-1}$cMpc) as that of QSO H II I-fronts. This is understandable because the same mechanism, i.e. density fluctuations and I-front shadows, is responsible for this dispersion.

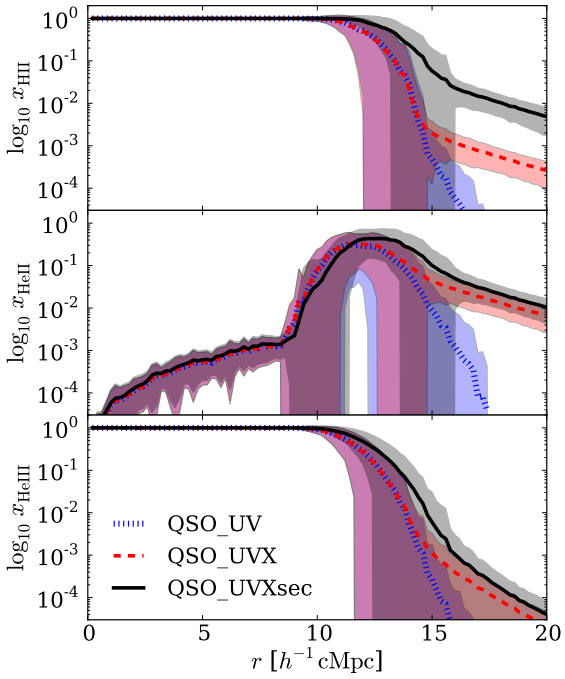

Figure 3.6: Spherically averaged profiles of H II (top), He II (middle), and He III (bottom) fractions around a QSO including only UV photons (QSO_UV, blue dotted), UV and X-ray photons (QSO_UVX, red dashed), UV and X-ray photon and secondary ionization (QSO_UVXsec, black solid). The shaded regions show the 1σ scatter of line-of-sight fluctuations.

3.3.2 Effect of X-rays and secondary ionization

Here we study the effect of X-rays and secondary ionization on the region ionized by the central QSO. This clarifies the physical origin of the extended tail of partial ionization in the QSO model seen in Figures 3.3 and 3.4.

Figure 3.6 shows the spherically averaged profiles of ionization fractions, $\langle x_i(r)\rangle_S$, together with the corresponding 1σ scatter, $\sigma_i^2(r) = \langle x_i^2(r)\rangle_S - \langle x_i(r)\rangle_S^2$, where $i =$ H II, He II, He III. The figure compares the three QSO only models with different RT physics, i.e. spectra restricted only to UV photons (QSO_UV), including UV+X-ray photons but neglecting the effect of secondary ionization (QSO_UVX), and including UV+X-ray photons and secondary ionization (QSO_UVXsec). The main impact of X-rays and secondary ionization on both hydrogen and helium is to produce an extended tail of partial ionization beyond the I-fronts.

In the case of hydrogen (the top panel of Figure 3.6), X-rays produce the partially ionized tail starting at $\sim 15h^{-1}$cMpc. The secondary ionization enhances the ionized fraction by an order of magnitude from $\langle x_{\rm H\,II}(r)\rangle_S \sim 10^{-3}$ (QSO_UVX) to $\langle x_{\rm H\,II}(r)\rangle_S \sim 10^{-2}$ (QSO_UVXsec) (see also [481]). The shaded regions indicate that the ionized fraction in the partially ionized tail has line-of-sight fluctuations about a factor of ~ 3 as a result of the density fluctuations[6].

Similarly, the partially ionized tails of He II and He III (the middle and bottom panels of Figure 3.6) have ionized fractions of $\langle x_{\rm He\,II}(r)\rangle_S \sim 10^{-2}$ and $\langle x_{\rm He\,III}(r)\rangle_S \sim 10^{-4} - 10^{-3}$. However, the impact of secondary ionization is less significant than on hydrogen reionization. The secondaries produce an enhancement of only a factor of ~ 2. This is because, due to its lower abundance, the probability that helium is collisionally ionized by fast electrons is a factor of $Y/X \simeq 0.09$ lower than hydrogen. Therefore, the secondary ionization (i.e. collisional ionization by fast electrons) affects little helium atoms.

Figure 3.7 shows the distributions of the H II and He III I-front radii (left panels) and their I-front thickness[7] (right panels) based on 100 random LOSs for the three QSO models discussed above. When neglecting secondary ionization, both distributions are weakly affected by X-rays, which is shown by a comparison between the two QSO models with (QSO_UVX, red dashed line) and without (QSO_UV, blue dotted line) X-rays. This is a direct consequence of the fact that most of the H I and He II-ionizing power is dominated by photons emitted near the ionization thresholds, i.e. the UV photons. X-ray photons alone do not affect the ionization state of the IGM near the I-fronts much, while the extended

[6] In addition, in all the QSO models the scatter is divergent (from $10h^{-1}$cMpc to $15h^{-1}$cMpc) near the I-front. This is because the I-front radii in each line-of-sights vary from $10h^{-1}$cMpc to $15h^{-1}$cMpc, the scatter around the spherical averaged profile becomes very large. Such divergent scatter is related to the line-of-sight fluctuations of the I-front positions. It is not a result of X-rays.

[7] The thickness is defined as the distance between the location where the ionized fraction is 90% and that at which it becomes 10%. As a first approximation, the thickness is about a mean free path, $\lambda_{\rm mfp}(\nu) = [n_{\rm H}(z)\sigma_{\rm H\,I}(\nu)]^{-1}$, of typical photons reaching the I-front (e.g. [419, 498, 236]). Although the thickness of the I-front is unresolved, i.e. $\lambda_{\rm mfp} < \Delta x_{\rm cell}$ where $\Delta x_{\rm cell} = 195h^{-1}$kpc is the cell size of our simulations, we use the above definition to support the claim that the presence of X-ray photons alone do not modify the thickness of the I-fronts.

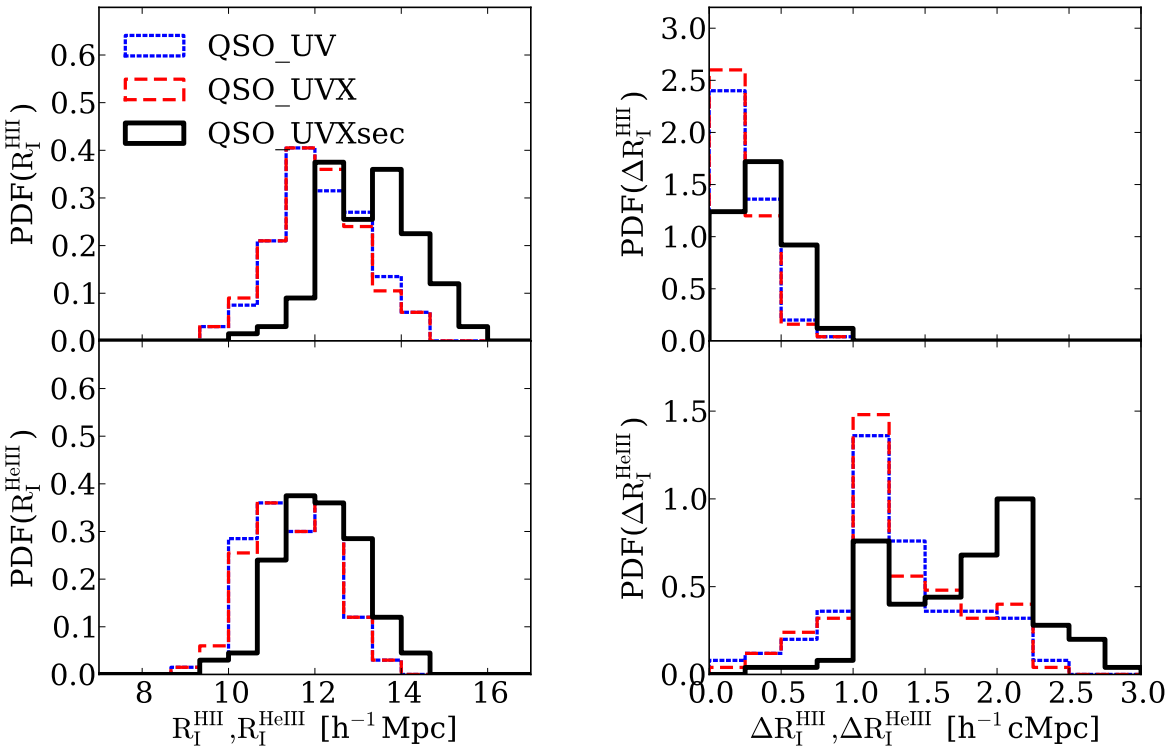

Figure 3.7: Distribution of the H II/He III I-front positions (left panels) and thickness (right panels) for the three QSO models including only UV photons (QSO_UV, blue dotted), UV and X-ray photons (QSO_UVX, red dashed), UV and X-ray photon and secondary ionization (QSO_UVXsec, black solid).

tail of partially ionized regions is indeed a strong feature of X-rays.

Including secondary ionization (QSO_UVXsec, black solid line), X-rays may, however, influence the distributions of the H II and He III I-fronts. First, the I-front positions appear to be pushed slightly outward by $\sim 1h^{-1}$cMpc (left panels). Second, the thickness is broadened by ~ 50 per cent. It could be a result of secondary ionization preferentially enhancing the ionization level in initially more neutral regions, which is causing a more gradual transition from ionized and neutral regions. The extent and the presence of such effects, though, seem to be related to the gas distribution (e.g. it is not observed in a homogeneous IGM simulation; Graziani, in prep.) and strongly depend on the version of the physics of secondary ionization implemented. Therefore, we restrict ourselves only to draw attention of the reader to these possible effects of secondary ionization near the I-fronts for future work.

3.3.3 H II and He II/He III regions in a reionization model driven both by galaxies and a QSO

We now analyse the more realistic case of switching a QSO on at $z = 10$ in a medium which has already been ionized by pre-existing galaxies, rather than being neutral as in § 3.3.1 and 3.3.2. This models the reionization scenarios driven by galaxies only and by both galaxies and a QSO. These models are designed to understand the role of a QSO in the galaxy overdensity during reionization.

We consider three different reionization histories by galaxies, which differ only for the total ionizing photon emissivity, namely models 2R, R, and 0.5R (Table 3.1). As a reference, Figure 3.8 shows the reionization histories of H II (red curves), He II (blue curves), and He III (green curves) fractions. The different linestyles correspond to the different galaxies only models. By $z = 10$, the pre-existing galaxies have driven H I reionization to the volume-averaged H II fractions of $\langle x_{\rm H\,II}\rangle_V \approx 28\%,\ 13\%,\ 6\%$ for 2R, R, 0.5R, respectively (similarly, the volume-averaged He II fractions are $\langle x_{\rm He\,II}\rangle_V \approx 26\%,\ 12\%,\ 5\%$). The ionization fractions increase by a factor of ~ 2 as the ionizing photon emissivities of the galaxies only models increase by the same factor. On the other hand, the contribution of the pre-existing galaxies to He II reionization is negligibly small.

In the following, we analyse how a QSO and galaxies together form the ionization state of the IGM in great detail.

Hydrogen

First of all, there is a significant impact of the pre-existing galaxies around a QSO on the morphology of the H II region. The top panels of Figure 3.9 show maps of H II fraction for both galaxies only and galaxies+QSO models. In the latter models, the central QSO is turned on at $z = 10$ in a medium that has been ionized by the pre-existing galaxies as described above.

Differently from the scenarios in § 3.3.1, H I reionization driven both by galaxies and a QSO (galaxies+QSO models) produces a morphology (size and shape) of the fully ionized

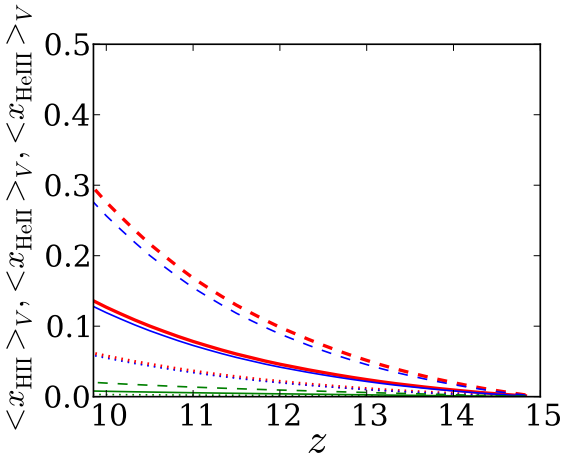

Figure 3.8: The simulated reionization history of the volume-averaged H II (red curves), He II (blue curves), and He III (green curves) fractions in the three galaxies only models: GAL(REF) (solid), GAL(2REF) (dashed), and GAL(0.5REF) (dotted).

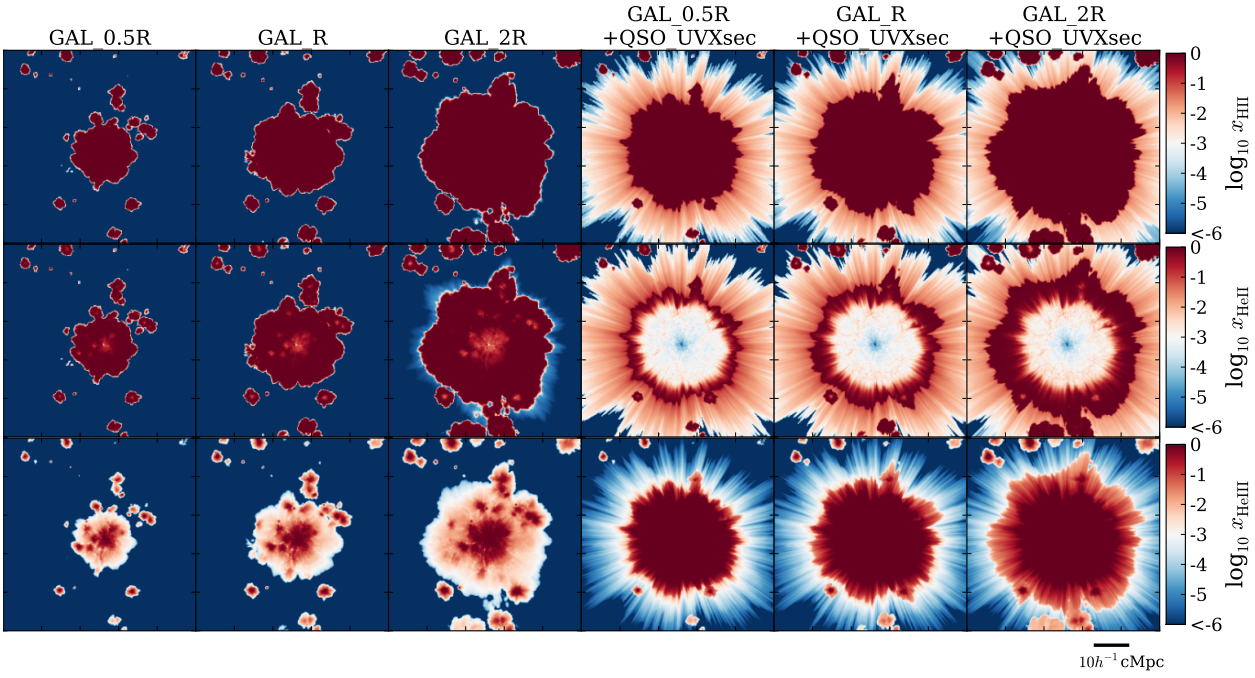

Figure 3.9: Maps of H II (top panels), He II (middle), and He III (bottom) fractions at $z = 10$ for galaxies only models GAL_0.5R, GAL_R, GAL_2R, and galaxies+QSO models GAL_0.5R+QSO_UVXsec, GAL_R+QSO_UVXsec, GAL_2R+QSO_UVXsec. In both sets of models, galaxies start to reionize the IGM at $z = 15$, while the QSO is turned on at $z = 10$ for a duration of 10^7 yr in the medium already reionized by the pre-existing galaxies. The maps have a side length of $50h^{-1}$cMpc and the width of a slice is $195h^{-1}$ckpc.

H II regions (with $x_{\rm H\,II} \gtrsim 0.9$) similar to one in the galaxies only environments (galaxies only models). This similarity is caused by the fact that a large quasi-spherical H II region can also be produced by the galaxies around a QSO. The galaxies produce a quasi-spherical H II region due to a combination of (i) the clustering of galaxies in the high density regions where QSOs preferentially reside, and (ii) the long timescale which galaxies have been shining for, allowing for the growth and overlap of many small H II regions.

At a fixed total emissivity of galaxies (i.e. runs with the same IDs 'GAL_R+...' etc), the QSO indeed contributes to produce a larger H II region (e.g. by $\sim 5h^{-1}$cMpc for GAL_R+QSO_UVXsec relative to GAL_R). However, once we compare two models with different total emissivities (e.g. GAL_2R vs. GAL_R+QSO_UVXsec), which is a reasonable variation given the uncertainty in escape fraction, the similarity in the size of H II region with and without a QSO is apparent. In fact, a QSO alone is capable of creating a H II region as large as an overlapped H II region around galaxies and vice versa. This means that there is an indistinguishability of sources responsible for a large H II region.

There are some distinctions between the galaxies only and galaxies+QSO models. The latter tends to produce slightly more spherical H II regions because the ionization is dominated by the central QSO during its activity (e.g. [103]). In addition, the prominent tail of partial ionization caused by the X-ray photons emitted by the QSO (see § 3.3.2) is still evident.

The left panels of Figure 3.10 show the distributions of the H II I-front radii for galaxies only and galaxies+QSO models. The radius obviously increases with the total ionizing photon emissivity. The dispersion of the H II I-front radii extends $> 5h^{-1}$ cMpc, which is larger than the dispersion for the QSO only models (see Figure 3.5). This is because the overlapped H II region produced by the surrounding galaxies is patchy. The distance to the I-front from the central galaxy/QSO varies significantly compared to one for more spherical H II regions found in the QSO only models. Therefore, when a QSO is surrounded by galaxies, the line-of-sight fluctuation of the H II I-front radii is caused both by (i) the patchiness of an overlapped H II region produced by the pre-existing galaxies and (ii) the density fluctuation as discussed in § 3.3.1.

Helium

The middle and bottom panels of Figure 3.9 show maps of He II and He III fractions for the galaxies only and galaxies+QSO models discussed above. Unlike the morphology of the H II regions, the morphology of the fully ionized He II/He III regions and the partially ionized tail have very distinctive features between galaxies only and galaxies+QSO models. In the galaxies only models He I reionization progresses simultaneously with H I reionization, and the morphology of He II and H II regions becomes very similar. As shown in the galaxies+QSO models, the impact of a QSO is then to push the He II I-front further outwards, increasing the extent of the pre-existed He II regions. At the same time, the hard spectrum of a QSO doubly ionizes helium behind the He II I-fronts of galaxies. This produces the He III region within the He II region, leaving only a shell of He II. Such shell is wider for larger ionizing photon emissivity of galaxies because the extent of the He II region created

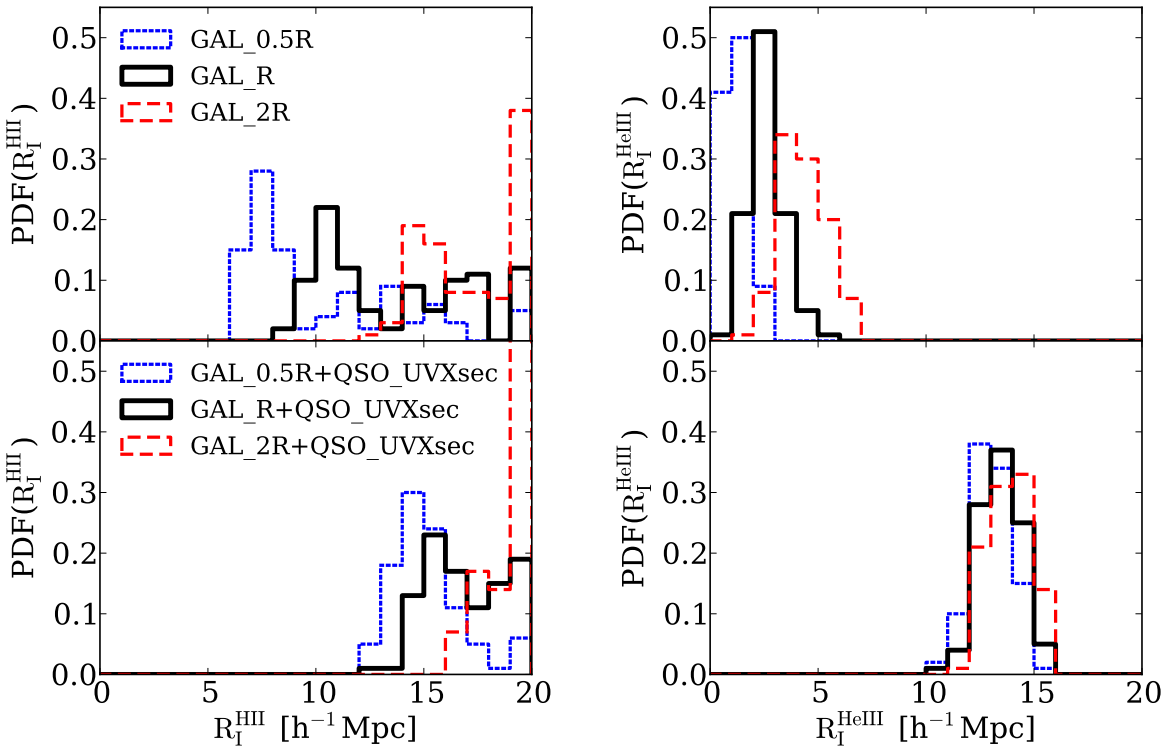

Figure 3.10: Distribution of the H II (right panels) and He III (left panels) I-front positions for galaxies only (top) and galaxies+QSO (bottom) models. The linestyle indicates the different source models of galaxies (i.e. different total ionizing photon emissivity): GAL_0.5R (blue dotted), GAL_R (black solid), and GAL_2R (red dashed).

by the pre-existing galaxies is larger. Because the QSO dominates the ionization of He II to He III, the He III region is highly spherical.

Figure 3.10 (right panels) shows the distributions of He III I-front radii for galaxies only and galaxies+QSO models. The central QSO pushes the He III I-front to $\sim 13h^{-1}$cMpc. As the QSO dominates the production of the He III region, the I-front radius is insensitive to the change of total ionizing emissivity of galaxies. In addition, the dispersion of the He III I-front is limited to only $\sim 5h^{-1}$cMpc, which is smaller than that of the H II I-front. This is because He II reionization is driven mostly by QSOs, the line-of-sight fluctuations of the He III I-fronts are caused only by the density fluctuations of the IGM (as discussed in § 3.3.1).

QSO-galaxy clustering

We highlight that the ratio of total cumulative ionizing photons emitted by galaxies and QSOs since the onset of reionization is the most important quantity for determining the ionization structure of the IGM during reionization. We elaborate on the role of galaxies around a QSO by the following simple illustrative argument.

The total number of ionizing photons of galaxies within comoving radius r around a QSO since the onset of reionization, $N_{ion}^{\rm GAL}(<r)$, can be estimated by

$$\begin{aligned} N_{ion}^{\rm GAL}(<r) &= \int_z^{z_0} dz' \left|\frac{dt}{dz'}\right| \dot{n}_{ion}(z') \int_0^r dr' 4\pi r'^2 [1+\xi_{qg}(r')], \\ &= \frac{4\pi}{3} n_{ion}^{\rm GAL}(z) r^3 \left[1 + \frac{4\pi}{3-\gamma}\left(\frac{r}{r_0}\right)^{-\gamma}\right], \end{aligned} \tag{3.6}$$

where $\xi_{qg}(r)$ is a QSO-galaxy correlation function, and $n_{ion}^{\rm GAL}(z) = \int_z^{z_0} dz' \, |dt/dz'| \, \dot{n}_{ion}(z')$, $|dt/dz| = 1/[H(z)(1+z)]$ is the integrated total comoving ionizing photon density emitted by galaxies until redshift z since the onset of reionization z_0. In the second equality, we assumed a power-law correlation function $\xi_{qg}(r) = (r/r_0)^{-\gamma}$ where r_0 is the correlation length and γ is the power-law index. By equating the number of ionizing photons and hydrogen atoms inside a radius $R_{\rm I,gal}^{\rm HII}$ (a.k.a. photon counting argument), the I-front radius for the pre-existing H II region around galaxies is estimated by $(4\pi/3)\bar{n}_{\rm H}(0)(R_{\rm I,gal}^{\rm HII})^3 = N_{ion}^{\rm GAL}(< R_{\rm I,gal}^{\rm HII})$ where $\bar{n}_{\rm H}(0)$ is the comoving hydrogen number density. By substituting equation (3.6), it simplifies to

$$R_{\rm I,gal}^{\rm HI} = r_0 \left[\frac{3-\gamma}{4\pi}\left(\frac{\bar{n}_{\rm H}(0)}{n_{ion}^{\rm GAL}(z)} - 1\right)\right]^{-1/\gamma}. \tag{3.7}$$

Using the best-fit power-law QSO-galaxy correlation function to our simulation, i.e. $r_0 = 18.6h^{-1}$cMpc and $\gamma = 1.6$, we can estimate the H II I-front radius as $R_{\rm I,gal}^{\rm HI} \approx 12h^{-1}$cMpc for our reference galaxies model (GAL_R), which is in reasonable agreement with the simulation (see Figure 3.10).

During the QSO activity, the QSO dominates the growth of cosmological I-front, starting with the pre-existing H II region of galaxies. Incorporating the effect of such pre-existing

H II region as an initial condition to the analytic equation of I-front (e.g. [398, 261]), the final I-front radius in a QSO environment, $R_{\rm I}^{\rm HI}$, is given by

$$\frac{4\pi}{3}(R_{\rm I}^{\rm HI})^3 - \frac{4\pi}{3}(R_{\rm I,gal}^{\rm HI})^3 = \frac{\dot{N}_{ion}^{\rm QSO}\bar{t}_{\rm rec,HII}}{\bar{n}_{\rm H}(0)}\left(1-e^{-t_Q/\bar{t}_{\rm rec,HII}}\right), \tag{3.8}$$

where $\bar{t}_{\rm rec,HII}$ is the recombination timescale of hydrogen. Assuming that the QSO lifetime is much shorter than the recombination timescale $t_Q \ll \bar{t}_{\rm rec,HII}$, the final I-front radius is given by

$$R_{\rm I}^{\rm HI} = R_{\rm I,gal}^{\rm HI}\left[1+\left(\frac{R_{\rm I,QSO}^{\rm HI}}{R_{\rm I,gal}^{\rm HI}}\right)^3\right]^{1/3}, \tag{3.9}$$

where $R_{\rm I,QSO}^{\rm HI} = \left[\frac{3\dot{N}_{ion}^{\rm QSO}t_Q}{4\pi\bar{n}_{\rm H}(0)}\right]^{1/3}$ is the comoving I-front radius of a QSO when it is turned on in isolation. Thus, using the above estimate of $R_{\rm I,gal}^{\rm HI}$ and the QSO model with $t_Q = 10^7$ yr, $R_{\rm I,QSO}^{\rm HI} \approx 11.8h^{-1}$cMpc, the analytic formula (3.9) gives $R_{\rm I}^{\rm HI} \approx 15h^{-1}$cMpc (again in reasonable agreement with simulation).

The fact that the simulated size of the H II region can be mostly explained by the above argument illustrates that the ionization structure of the IGM largely depends on (i) the total cumulative ionizing photon emissivity from galaxies, $n_{ion}^{\rm GAL}(z)$, (ii) the QSO-galaxy correlation function, $\xi_{qg}(r)$, and (iii) the ionizing photons emitted by a QSO, $\dot{N}_{ion}^{\rm QSO}t_Q$. The final size of the H II region (equation 3.9) depends on $(R_{\rm I,QSO}^{\rm HI}/R_{\rm I,gal}^{\rm HI})^3 \propto \dot{N}_{ion}^{\rm QSO}t_Q/N_{ion}^{\rm GAL}$. In other words, the above argument, although highly simplified, demonstrates that *the ratio of the integrated QSO luminosity, $\dot{N}_{ion}^{\rm QSO}t_Q$, to the integrated total ionizing photons of galaxies around a QSO, $N_{ion}^{\rm GAL}(<r)$, since the onset of reionization* plays a central role in determining the size of H II regions in QSO environments. If the total photons from galaxies since the onset of reionization exceed or are comparable to the photons from a QSO during its activity, the distinction of the morphology of H II region between the ones by galaxies and a QSO becomes vague. On the other hand, if a high-redshift QSO could shine in the environment that the ionizing photon from galaxies is sub-dominant, then a QSO alone produces a large spherical H II region.

3.3.4 Dependence on the properties of a QSO

So far we have focused on the role of galaxies and a QSO in shaping the ionization state of the IGM. However, the QSO was assumed to have a power-law spectrum and a single episode of QSO activity. In the following, we analyse the RT simulation suite to understand the dependence of the impact of QSOs on the spectral properties and the QSO duty cycle.

Effect of an obscured QSO

The effect of an obscured QSO demonstrates how the X-ray sources without UV emission influence the ionization state of the IGM. To isolate the effect of the different spectrum

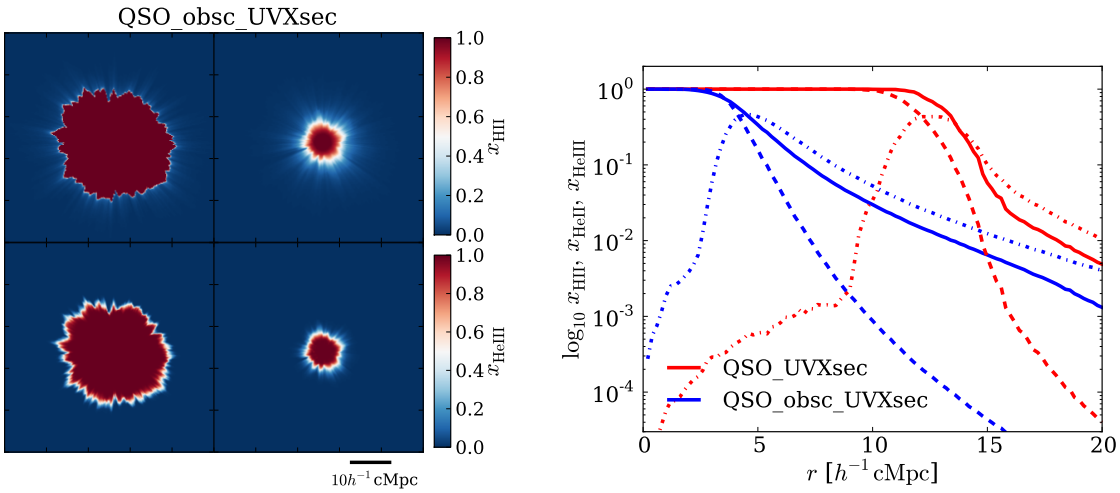

Figure 3.11: (Right) The spherically averaged profiles of H II (solid), He II (dotted dash), and He III (dashed) fractions for QSO (QSO_UVXsec, red curves) and obscured QSO (QSO_obsc_UVXsec, blue curves) models. (Left) The linear-scale maps of H II (top) and He III (bottom) fractions at $z = 10$ of the QSO (QSO_UVXsec, left column) and obscured QSO (QSO_obsc_UVXsec, right column) models. The linear scale is chosen to emphasize the ionization structure near the I-fronts.

of a QSO, we first use the two QSO only models only changing the type of spectrum (QSO_UVXsec and QSO_obsc_UVXsec runs). Figure 3.11 shows the maps of (unobscured) QSO only model (QSO_UVXsec) and obscured QSO only model (QSO_obsc_UVXsec) on left, and the spherically averaged profiles of H I, He II, and He III fractions on right. A first obvious difference is that the obscured QSO model shows smaller H II and He III regions. This is because the ionizing photons between 13.6 eV and 100 eV are absorbed by the neutral hydrogen gas inside a QSO host galaxy (see Figure 3.2). For the obscured QSO model, the photoionization is driven only by the soft X-ray photons. A second difference is in the shape of I-fronts. The lack of UV photons in the obscured QSO model and the large mean free path of X-ray photons create much smoother I-fronts than those of the unobscured QSO model. Note that the presence of UV photons in a QSO spectrum (QSO_UVXsec) is already sufficient to produce a sharp I-front. The effect of an obscured QSO is to produce a broad I-front because it satisfies the two conditions: (i) the presence of X-ray photons and (ii) the lack of UV photons in the spectrum.

Similarly to § 3.3.3, the surrounding galaxies play an important role in how an obscured QSO impacts the ionization state of the IGM. Figure 3.12 plots the linear-scale maps of H II and He III fractions for galaxies only model (GAL_R), galaxies+(unobscured) QSO model (GAL_R+QSO_UVXsec) galaxies+obscured QSO model (GAL_R+QSO_obsc_UVXsec) at the snapshots corresponding to the two QSO lifetimes at $t_Q = 10^7$ yr and 8×10^7 yr.

The morphology of H II regions (top panels of Figure 3.12) shows that, unlike the obscured QSO only model, in galaxies+QSO models the obscured QSO has little impact

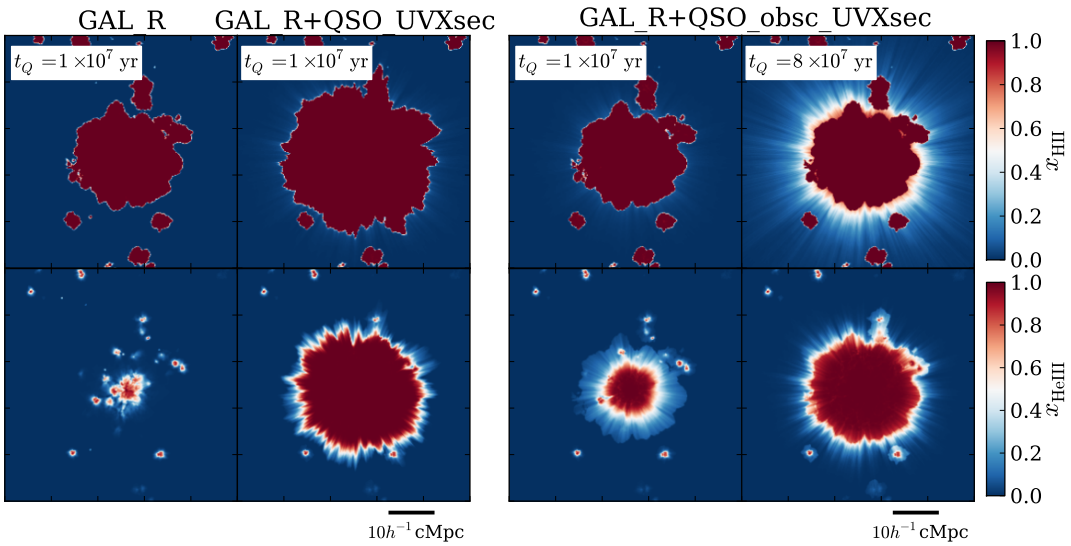

Figure 3.12: The linear-scale maps of H II (top) and He III (bottom) fractions for galaxies only model (GAL_R), galaxies+QSO model (GAL_R+QSO_UVXsec), galaxies+obscured QSO model (GAL_R+QSO_obsc_UVXsec) at $t_Q = 10^7$yr (second right column) and $t_Q = 8 \times 10^7$yr (rightmost column). The QSO is turned on at $z = 10$ in the medium already reionized by galaxies. The length of x- and y-axes is $50h^{-1}$cMpc and the width of the slice is $195h^{-1}$ckpc. The linear scale was chosen to show structure near the I-front more clearly.

on the shape and size of the H II region at $t_Q = 10^7$ yr. This is because the surrounding galaxies provide another sources of UV photons, the galaxies still dominate the growth of the H II I-front when the total ionizing soft X-ray photons from the obscured QSO is small. At $t_Q = 8 \times 10^7$ yr (the rightmost panel of Figure 3.12), the ionizing photons from the obscured QSO start to overtake that of galaxies. As the growth of the H II I-front is now dominated by the obscured QSO, it imprints a unique and smoother I-front around the central H II region. Because of this, galaxies+obscured QSO model can imprint a more distinctive feature in the morphology of the H II region than that of galaxies+(unobscured) QSO model (cf. § 3.3.3).

On the other hand, because He II reionization is driven mostly by QSOs, the morphology of He III regions (the bottom panels of Figure 3.12) is directly affected by the type of QSOs. The obscured QSO produces smoother He III I-front than that of an unobscured QSO model. The effect and physical mechanism of shaping the QSO He III region are therefore identical to QSO only models.

In summary, the impact of an obscured QSO is to produce a very broad I-front. The effect is noticeable only if the amount of soft X-ray photons from obscured QSOs exceeds UV photons from galaxies. In other words, the net composite spectrum of the collective radiation field from galaxies and QSOs must have dominant X-ray emissions in order for obscured QSOs to produce distinctive H II regions in their local environments.

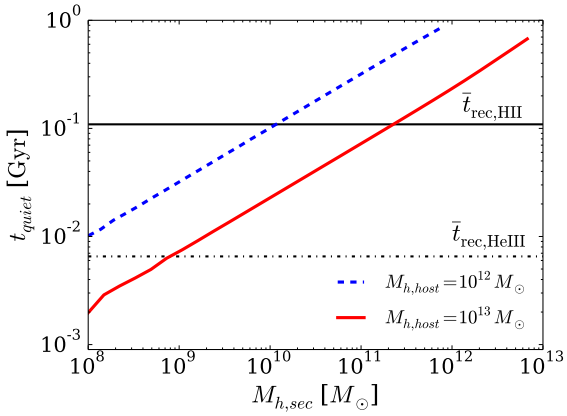

Figure 3.13: The halo merger timescale and recombination timescale as a function of secondary halo mass merging to a QSO-host halo. The horizontal lines are the recombination timescale of hydrogen and helium. The figure indicates that the QSO activity driven by the mergers with the timescale $t_{quiet} < \bar{t}_{\rm rec,HII}$ ($t_{quiet} > \bar{t}_{\rm rec,HII}$) can (cannot) sustains the H II region. The same argument also holds for He III region.

Effect of QSO duty cycle

We investigate the impact of QSO duty cycle on reionization. We use an analytic argument in this section. During the course of the work, we have also run a RT simulation explicitly including the effect of QSO duty cycle. Because the previous activity of a QSO has trivially increased the size of ionized regions and there is a large uncertainty in the triggering mechanism of QSO activity at high redshift, we think that a simple analytic argument, which allows us to scan through a large parameter space, is more illustrative as a first step. Of course, a future quantitative study should use RT simulations with QSO duty cycle.

The build-up of SMBH driving the QSO activity requires $t = t_{\rm Edd}\ln(M_{\rm BH}/M_{seed}) \sim 5\times10^7$ yr with the seed mass of $M_{seed} = 10^{2-5}$ M$_\odot$ and $t_{\rm Edd} = 4\times10^7$ yr even with Eddington limit accretion rate. A QSO-host galaxy may thus have experienced multiple episodes of QSO activity (as our fiducial QSO lifetime $t_Q = 10^7$ yr is smaller than the timescale to build up a SMBH). Assuming that the QSO activity is triggered by merger events of haloes (e.g. [195]), the duration of the QSO-quiet phase, t_{quiet}, is expected to be of order of the halo merger timescale $t_{quiet} \approx t_{merge}$.

To get an idea of how the multiple QSO activities impact the reionization, we introduce a timescale-based argument. In this argument, the QSO-quiet phase timescale is compared with the recombination timescale, $\bar{t}_{rec,i}$, of i-th species i = H II, He III. If the QSO-quiet phase timescale is shorter than the recombination rate, $t_{quiet} < \bar{t}_{rec,i}$, the long recombination rate allows the IGM to remain ionized (fossil H II region) during the QSO-quiet phase ([156, 444]). Then, the H II region in the QSO environment can regrow at the next active QSO phase as if the QSO has remained turn on all the time. On the other hand, when

the QSO-quiet timescale is longer than the recombination timescale, $t_{quiet} > \bar{t}_{rec,i}$ the IGM recombines and becomes neutral before the next QSO phase. Hence, each QSO episode must restart ionizing the IGM from neutral because the QSO-quiet phase is very long so that the medium recombines before the next QSO activity. In this limit, only the single phase of QSO activity at a redshift of interest contributes to reionization.

Armed with the above timescale argument, we estimate the halo merger timescale, hence the QSO-quiet timescale, from the extend Press-Schechter approach ([242]) as an inverse of the halo merger rate at $z = 10$. Figure 3.13 shows the QSO-quiet phase timescale of host halo mass $M_{h,host}$ (10^{13} $M_\odot$ (red solid) and 10^{12} $M_\odot$ (blue dashed)) merging with secondary halo mass $M_{h,sec}$. The recombination timescales of hydrogen (H II $\rightarrow$ H I) and helium (He III $\rightarrow$ He II) at mean density, $\bar{t}_{\rm rec,HII} = [\alpha_{\rm B,HII}\bar{n}_{\rm H}(z)C]^{-1}$ and $\bar{t}_{\rm rec,HeIII} = [\alpha_{\rm B,HeIII}\bar{n}_{\rm He}(z)C]^{-1}$ with a clumping factor $C = 3$,[8] are shown as the horizontal black lines.

Figure 3.13 shows the duration of QSO-quiet phase (halo merger timescale) of the two QSO-host halo masses as a function of secondary halo mass and the hydrogen and helium recombination timescales. The comparison between the two timescales in Figure 3.13 shows that to keep the hydrogen reionized ($t_{quiet} < \bar{t}_{rec,\rm HII}$) the QSO need to be fed by minor mergers with halo of mass $M_{h,sec} \lesssim 10^{10}(10^{11})$ $M_\odot$ for host halo of mass $M_{h,host} = 10^{12}(10^{13})$ $M_\odot$; otherwise, the QSO H II region can recombine before the next phase of QSO activity. Similar is true for He III region. Because of the faster recombination timescale of He III, the possible onset of He II reionization around a high-z QSO ([40]) should be fed even more vigorously than hydrogen to maintain the QSO He III region. Thus, the impact of QSO duty cycle is controlled by the ratio between the timescale of QSO activity and the recombination timescale.

3.4 Thermal state of the IGM

Our results on the thermal structure of the IGM in the QSO environment are presented and analysed in this section. Briefly, the concerted progress of hydrogen and helium reionization driven by galaxies and QSOs impacts non-linearly the thermal state of the IGM. We present our results with increasing sophistication to highlight the effect of photoionization heating across the I-fronts (§ 3.4.1), pre-heating by X-rays (§ 3.4.2), and the additional complexity when both galaxies and QSOs are taken into account (§ 3.4.3).

3.4.1 Thermal state of the IGM in a galaxy- or QSO-driven reionization model

We first present the simplest case where the heating is due to photoionization driven only by galaxies or a QSO turned on at $z = 10$ in fully neutral initial conditions. This aims at

[8] We assumed $C = 3$ for simplicity. Any change in the clumping factor linearly scales the timescale shown in Figure 3.13 up and down.

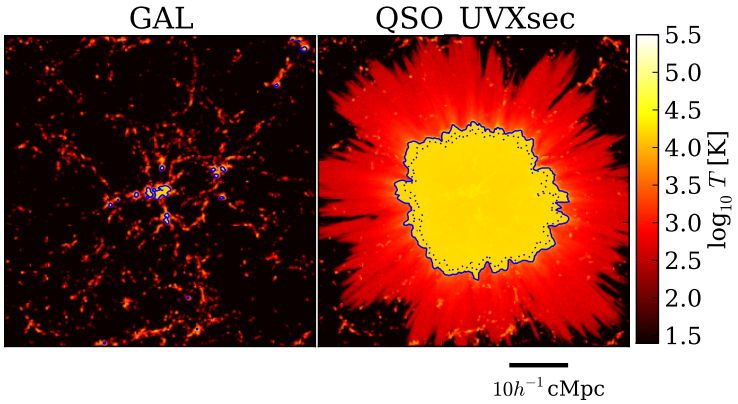

Figure 3.14: Maps of temperature at $z = 10$ for galaxies only model, GAL (left), and QSO only model, QSO_UVXsec (right). In both models the sources are turned on in a fully neutral initial condition at $z = 10$ and shine for 10^7 yr. The maps have a side length of $50h^{-1}$cMpc and the width of the slice is $195h^{-1}$ckpc. The H II (He III) I-front is shown as blue solid (dotted) contour.

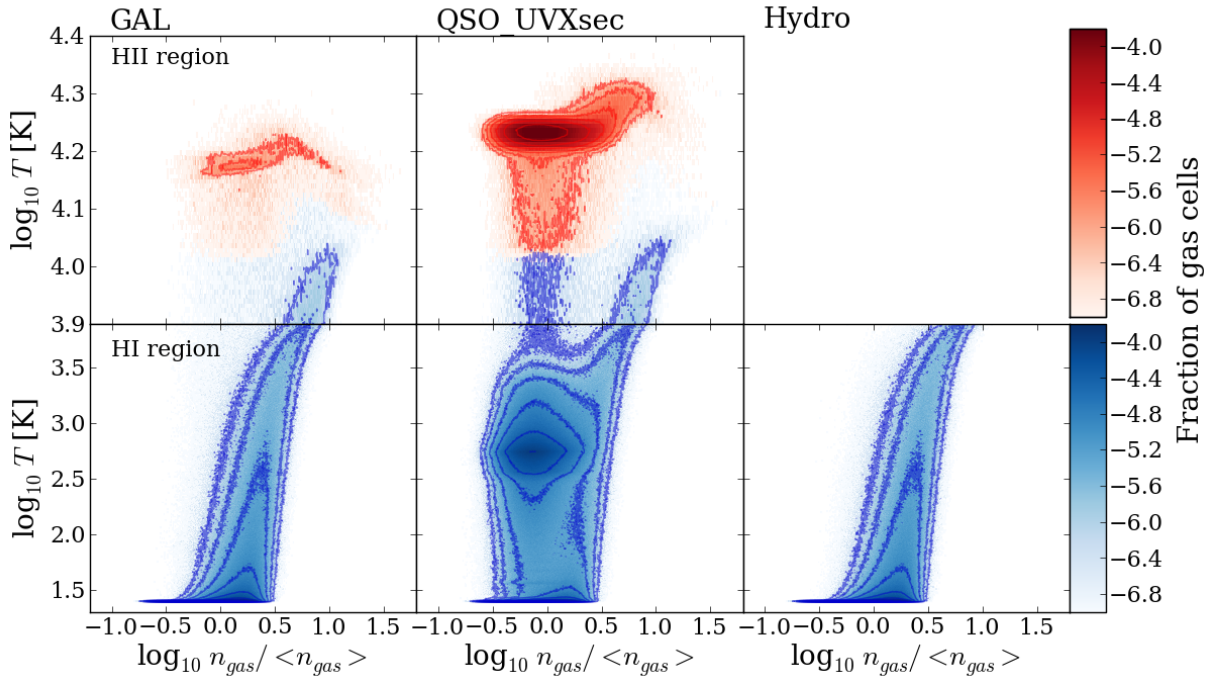

Figure 3.15: Temperature-density diagram at $z = 10$ for models GAL (left), QSO_UVXsec (middle), and for the hydrodynamic simulations without RT (right). The red (blue) colour indicates the H II (H I) regions defined as cells with $x_{\rm H\,II} \geq 0.5 (< 0.5)$. The contours are overlaid as a guide. To emphasize the temperature-density relation both in H II and H I regions, two different y-axis scales are used.

understanding the thermal structure of the IGM in galaxy- and QSO-driven reionization models. Figure 3.14 shows temperature maps for the galaxies only model (GAL, left) and the QSO only model (QSO_UVXsec, right). The contours refer to the position of the H II (blue solid) and He III (blue dotted) I-fronts. In both models, the filamentary thermal structure in the H I region is a result of adiabatic compression and hydrodynamic shocks along filaments of cosmic web. The passage of I-fronts heats up the gas inside the H II region by photoionization to $\gtrsim 10^4$ K, and therefore the thermal structure is well traced by the position of I-fronts. Similarly to the ionization structure discussed in § 3.3.1, a large central heated region is present around a QSO, whereas galaxies produce small and patchy heated regions. X-rays from a QSO pre-heat the gas ahead of the sharp I-front, which will be discussed in detail in § 3.4.2.

Figure 3.15 shows the temperature-density diagram for the two source models discussed above, as well as for the hydrodynamical simulation without RT. The colour indicates H II (H I) regions where the local H II fraction is $x_{\rm H\,II} \geq 0.5 (< 0.5)$. By comparing the galaxies only (GAL, left) and QSO only (QSO_UVXsec, middle) models, we can see that the temperature of the QSO H II region (~ 17000 K) is $\sim 16\%$ higher than the one of galaxies' H II regions (~ 14700 K). This is because of the harder spectrum, the central QSO heats up the gas more effectively per H I and He I ionization (see later section 'photoionization heating') and provides an additional heating by He II photoionization. The temperature of the H II region is nearly isothermal in both models. Because the gas is ionized within a short period of time (10^7 yr in these models), the most of photoionized gas is heated simultaneously and still remains isothermal. At longer timescale, the isothermality is broken because the adiabatic (recombination) cooling preferentially cools low (high) density regions of the IGM. In the temperature-density relation of the H I region, comparing the leftmost and rightmost panels (GAL vs. Hydro) of Figure 3.15, it is clear that the temperature in the galaxies only model is solely determined by hydrodynamic processes. This is because the UV photons from galaxies cannot penetrate into neutral gas, only mechanisms that can control the thermal state are hydrodynamical processes. On the other hand, X-rays from a QSO penetrate deep into neutral regions ahead of the H II I-front. This pre-heats the neutral gas, which is partially ionized to $x_{\rm H\,II} \sim 10^{-2}$, to $\sim 10^3$ K (see also § 3.4.2).

Photoionization heating

The degree of the IGM temperature jump across the passage of I-fronts depends on the spectral shape of the sources, and is insensitive to their ionizing photon luminosities. Photoionization by UV and X-ray radiations heats up the gas, and the thermal energy injected by photoionization is[9] $G_i = \int_{\nu_i}^{\infty} \frac{4\pi J_\nu}{h\nu} \sigma_i(\nu)(h\nu - h\nu_i) d\nu$ and the photoionization rate is $\Gamma_i = \int_{\nu_i}^{\infty} \frac{4\pi J_\nu}{h\nu} \sigma_i(\nu) d\nu$ where J_ν [$= L_\nu/(4\pi r)^2$ in the optical thin limit] is the specific intensity at a radius r, and $h\nu - h\nu_i$ is the excess energy of photoionization above the ionization threshold and the index run for each species, $i = {\rm H\,I, He\,I, He\,II}$. The heat gain

[9] We ignore the impact of secondary ionization for this estimate, although it is included in the simulations.

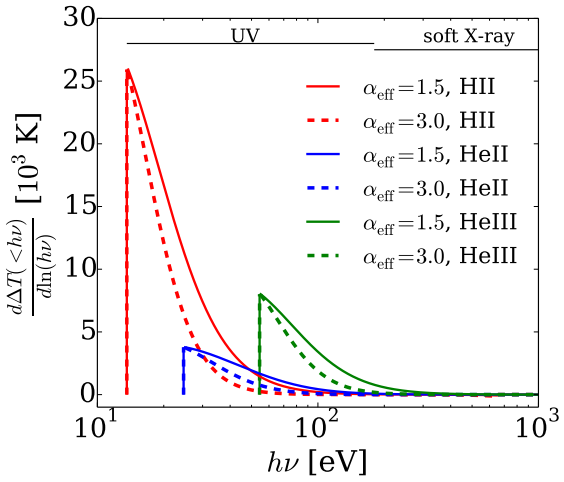

Figure 3.16: The temperature gain per photon energy in logarithmic bins. The red, blue, and green lines indicate the photoionization heating across H II, He II, and He III I-fronts, respectively. The linestyles indicate the type of spectrum: $\alpha_{\rm eff} = 1.5$ for galaxy-type (dashed), $\alpha_{\rm eff} = 3.0$ for QSO-type (solid). The figure shows that the most contribution to photoionization heating is from UV photons.

per ionization across an I-front is (e.g. [1])

$$\langle E_i \rangle = G_i/\Gamma_i \approx \frac{h\nu_i}{2+\alpha_{\rm eff}}, \tag{3.10}$$

where $\alpha_{\rm eff}$ is the effective spectral index at the position of I-front. In the optical thin limit, it is identical to the spectral index of source $\alpha_{\rm eff} = \alpha_G$ or α_Q. If the spectral hardening across the I-fronts occurs, the effective index approaches to harder spectral value $\alpha_{\rm eff} < \alpha_G$ or α_Q.

From the energy conservation across the I-front, we obtain $(3/2)k_B n_{gas}^{\rm after}\Delta T \approx n_i^{\rm before}\langle E_i\rangle$ where ΔT is the temperature jump across a I-front, and $n_{gas}^{\rm after}$ ($n_i^{\rm before}$) is the number density of gas of i-th species after (before) the passage of the I-front. Thus for normal ordering of I-fronts (H II I-front precedes ahead of He II/He III I-fronts), the temperature jump across H II I-front with incident ionizing photons with energy $< h\nu$ is given by

$$\Delta T_{\rm HII}(<h\nu) = \frac{2}{3k_B(2+Y/X)} \frac{\displaystyle\int_{\nu_i}^{\nu} (\nu/\nu_{\rm HI})^{-\alpha_{\rm eff}} \sigma_{\rm HI}(\nu)(h\nu - h\nu_{\rm HI})\frac{d\nu}{h\nu}}{\displaystyle\int_{\nu_i}^{\nu} (\nu/\nu_{\rm HI})^{-\alpha_{\rm eff}} \sigma_{\rm HI}(\nu)\frac{d\nu}{h\nu}}. \tag{3.11}$$

Similar formulae hold for He II and He III I-fronts, but with different coefficients.

The differential temperature jump across an I-front per incident photon energy, $d\Delta T(<h\nu)/d\ln(h\nu)$, is shown in Figure 3.16. It shows that UV photons dominate the photoionization heating across I-fronts. Additional soft X-ray contribute little to the net heating across I-fronts.

The total temperature jumps ($h\nu \to \infty$) across H II, He II, and He III I-fronts are then given by[10]

$$\Delta T_{\rm HII} = \frac{2}{3k_B}\frac{\langle E_{\rm HI}\rangle}{2+Y/X} \approx 5.0\times10^4(2+\alpha_{\rm eff})^{-1}\ {\rm K}, \tag{3.12}$$

$$\Delta T_{\rm HeII} = \frac{2}{3k_B}\frac{\langle E_{\rm HeI}\rangle}{2(X/Y)+2} \approx 7.6\times10^3(1+\alpha_{\rm eff})^{-1}\ {\rm K}, \tag{3.13}$$

$$\Delta T_{\rm HeIII} = \frac{2}{3k_B}\frac{\langle E_{\rm HeII}\rangle}{2(X/Y)+3} \approx 1.6\times10^4(2+\alpha_{\rm eff})^{-1}\ {\rm K}. \tag{3.14}$$

Thus the temperature jump after the passage of all the I-fronts is $\approx 2.2\times10^4$ K for QSO-type spectrum $\alpha_{\rm eff}=1.5$, whereas $\approx 1.5\times10^4(1.2\times10^4)$ K for galaxy-type spectrum $\alpha_{\rm eff}=3$ (without He II reionization). In this estimate, QSO-type spectrum is expected to heat up the gas more effectively by $\sim 80\%$ than galaxy-type spectrum. This is somewhat larger than the difference found in our simulations. As shown in § 3.4.2, neglecting secondary ionization, the temperature inside the QSO H II region becomes ~ 20000 K. In addition, the cooling is neglected in the simplified argument. Taking into account these two factors, the above estimate explains, at least to the accuracy expected from this simplified argument, the difference in the temperature seen in the simulations.

The above simplified argument demonstrates the physical origin of the temperature jump across the H II, He II, and He III I-fronts, i.e. the energy conservation of spectral-weighted excess energy of photo-electrons across the I-fronts and the change of gas particle number. A harder spectrum more efficiently heats up the gas per ionization because the fraction of higher energy photons relative to photons near the ionization threshold is higher for a smaller spectral index α (e.g. [1]).

3.4.2 Effect of X-rays and secondary ionization

Figure 3.17 shows the spherically averaged profiles of temperature for three QSO only models including only UV photons (QSO_UV), UV and X-ray photons but neglecting secondary ionization (QSO_UVX), UV and X-ray photons and secondary ionization (QSO_UVXsec). At the tail of partially ionized region extending from $\gtrsim 15h^{-1}$ cMpc, X-ray photons from a QSO pre-heat the gas to $\sim 10^3$ K. Including secondary ionization slightly lowers the temperature by $\sim 10\%$. This is expected because secondary ionization diverts some of the total excess energy of fast photo-electrons into collisional ionization of hydrogen and helium rather than heating.

Inside the H II region at $< 10h^{-1}$ cMpc, X-ray photons does not affect the temperature (comparing QSO_UV with QSO_UVX). This is understandable because the temperature in the fully ionized regions depends mostly on the UV photons through the photoionization

[10]For this estimate, we have approximated the photoionization cross sections of H I, He I, and He II as power-laws, $\sigma_{\rm HI}\approx 6.3\times10^{-18}(\nu/\nu_{\rm HI})^{-3}$, $\sigma_{\rm HeI}\approx 7.8\times10^{-18}(\nu/\nu_{\rm HeI})^{-2}$, $\sigma_{\rm HeII}\approx 1.6\times10^{-18}(\nu/\nu_{\rm HeII})^{-3}$, respectively ([128]).

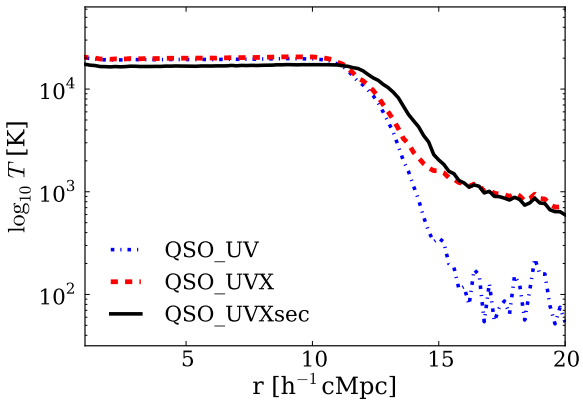

Figure 3.17: Spherically averaged temperature profiles around a QSO for three QSO only models including only UV photons (QSO_UV, blue dotted), UV and X-ray photons (QSO_UVX, red dashed), UV and X-ray photon and secondary ionization (QSO_UVXsec, black solid).

heating at the I-fronts (as shown in § 3.4.1). Including secondary ionization, however, lowers the post-ionized temperature by $\sim 18\%$ from ~ 20000 K to ~ 17000 K.

Near the I-fronts at $\sim 10 - 15h^{-1}$cMpc, the QSO model with secondary ionization shows higher temperature than the model without secondaries. This is related to the fact that our secondary ionization model pushes the I-front slightly outwards. This means that the location of the photoionization heating by UV photons at the I-front is displaced when including secondary ionization. Therefore, secondary ionization indirectly increases the radius of the photoheated region.

In summary, X-ray photons and secondary ionization impact the thermal state in two ways: (i) X-ray photons heat the gas beyond the I-front because of their larger mean free path, and (ii) secondary ionization lowers the net heating efficiency, resulting in slightly lower temperature than a case when it is ignored.

3.4.3 The thermal state of the IGM in a reionization model driven both by galaxies and a QSO

If a QSO turns on in a medium previously ionized by pre-existing galaxies, the thermal state of the IGM is determined by the collective impacts of both galaxies and the central QSO. Here we discuss our most realistic case for reionization models driven both by galaxies and a QSO.

Figure 3.18 shows temperature maps for galaxies only models (GAL_0.5R, GAL_R, GAL_2R) and galaxies+QSO models (GAL_0.5R+QSO_UVXsec, GAL_R+QSO_UVXsec, GAL_2R+QSO_UVXsec), in which the QSO is turned on in a medium previously ionized by pre-existing galaxies. The blue solid (dotted) contours indicate the location of H II (He III) I-fronts. As already noticed in Figure 3.14, the thermal structure is traced by the

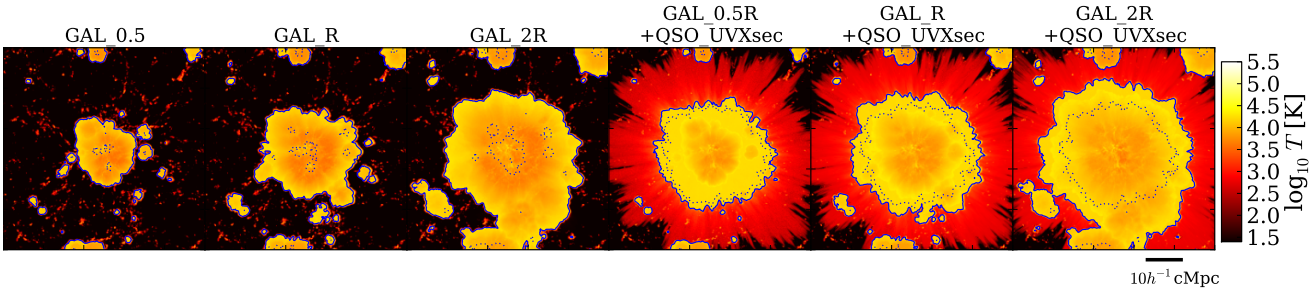

Figure 3.18: Maps of temperature at $z = 10$ for galaxies only models GAL_0.5R, GAL_R, GAL_2R, and galaxies+QSO models, GAL_0.5R+QSO_UVXsec, GAL_R+QSO_UVXsec, GAL_2R+QSO_UVXsec. In both sets of models galaxies start to reionize the IGM at $z = 15$, while the QSO is turned on at $z = 10$ for a duration of 10^7 yr in the medium already reionized by the pre-existing galaxies. The maps have a side length of $50h^{-1}$cMpc and the width of the slice is $195h^{-1}$ckpc. The H II (He III) I-front is shown as blue solid (dotted) contour.

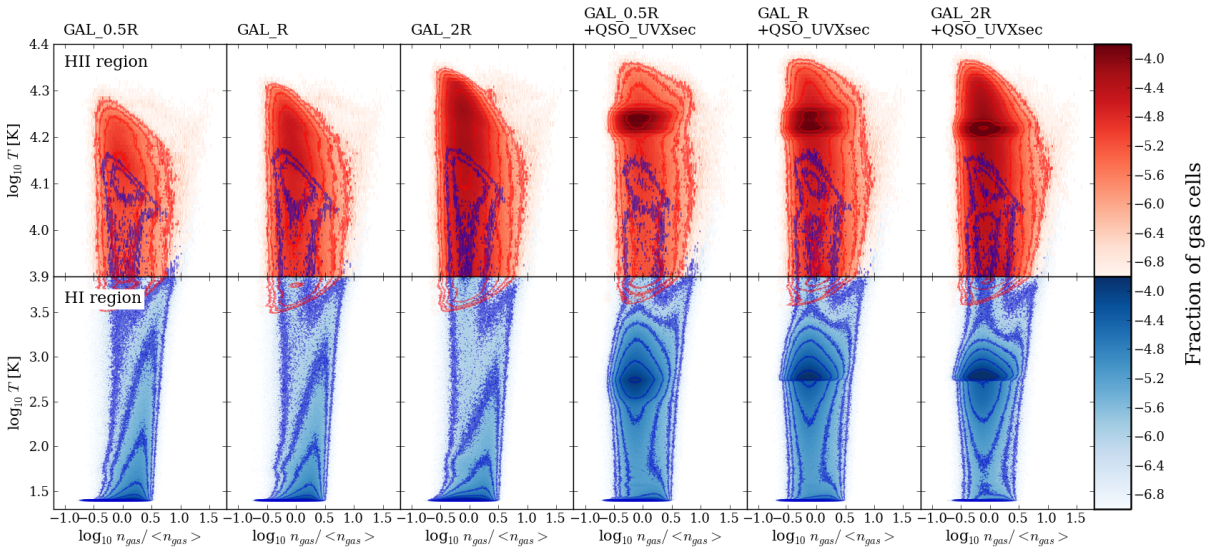

Figure 3.19: Temperature-density diagram at $z = 10$ for galaxies only models GAL_0.5R, GAL_R, GAL_2R and galaxies+QSO models GAL_0.5R+QSO_UVXsec, GAL_R+QSO_UVXsec, GAL_2R+QSO_UVXsec. The red (blue) colour indicates the H II (H I) regions defined as cells with $x_{\rm H\,II} \geq 0.5 (< 0.5)$. The contours are overlaid as a guide.

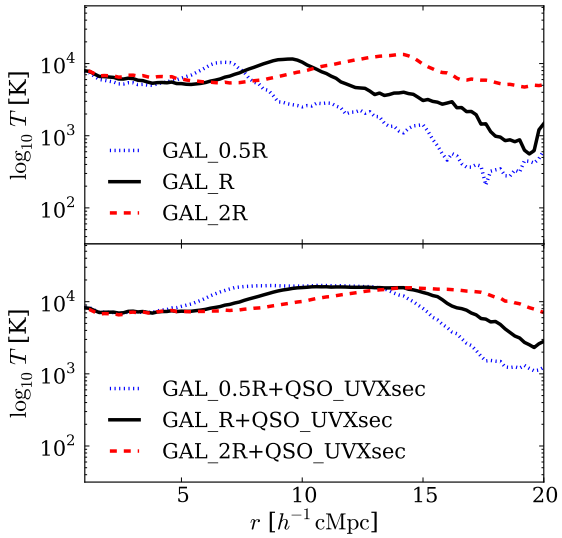

Figure 3.20: Spherically averaged temperature profiles of galaxies only (top) and galaxies+QSO (bottom) models varying the total ionizing photon emissivity with full RT physics. In the top panel, blue dotted, black solid, and red dashed lines indicate models GAL_0.5R, GAL_R, GAL_2R. In the bottom panel, blue dotted, black solid, and red dashed lines indicate models GAL_0.5R+QSO_UVXsec, GAL_R+QSO_UVXsec, GAL_2R+QSO_UVXsec.

position of the H I I-front due to the photoionization heating of hydrogen. Differently from a case in which reionization is driven by a QSO only (§ 3.4.1), pre-existing galaxies drive the H II I-fronts further ahead of the He III I-fronts, which are driven mainly by the QSO. Thus, a large fraction of the thermal structure inside the H II region is determined by the pre-existing galaxies instead of the central QSO. The effect is more obvious as the total ionizing photon emissivity from galaxies increases. Similarly to what already noticed for the H II regions (§ 3.3.3), except for the pre-heating by X-rays from a QSO ahead of the H II I-front, the morphology and extent of the gas heated at $\sim 10^4$ K is very similar in both galaxies only and galaxies+QSO models, as it is mainly determined by the UV photons emitted by galaxies.

We examine the thermal structure in more detail, focusing on the effect of galaxies and the additional heating by a QSO.

Heating by galaxies

Galaxy-type spectrum contributes to the thermal structure of the IGM primarily by H I and He I photoionization heating. In Figure 3.19 the temperature-density diagram of H II (red) and H I (blue) regions shows that in galaxies only models (left three panels) the maximum temperature reached ($\sim 2 \times 10^4$ K) weakly depends on the change in total ionizing photon

emissivity, whereas a larger fraction of gas is heated near to the maximum temperature with increasing total emissivity. This is because, for a fixed source spectrum, the heating rate per ionization is fixed (§ 3.4.1), and thus the heating efficiency remains the same regardless of the change of the total ionizing luminosity.

When the galaxies ionize the IGM gradually from $z = 15$ to 10, the temperature distribution in the H II regions is no longer isothermal and a clear spread is visible [in contrast to the case in which galaxies or a QSO shines in a fully neutral region for a short period of time (as discussed in § 3.4.1)]. This is because the gas that is engulfed by the I-front earlier (inner parts of the H II region) has more time to cool. Figure 3.20 (top panel) plots the spherically averaged temperature profiles in galaxies only models. Temperature in the inner H II region is lower than the gas temperature immediately behind the I-front[11]. As the cooling takes place differently depending on the local gas density inside the H II region, the temperature spread is large when the period of reionization is extended and driven gradually by galaxies.

On the other hand, the thermal structure in H I regions is nearly identical in all galxies only models. Similarly to § 3.4.1, UV photons from galaxies do not influence the temperature ouside H II regions. Hydrodynamical processes are responsible for controlling the temperature in the neutral gas.

Collective heating by a QSO and galaxies

QSO-type spectrum contributes to the thermal structure by the H I and He I photoionization heating as well as the He III photoheating due to its harder spectrum. In addition, X-ray preheating ahead of the H II I-front is a distinctive heating mechanism by QSO-type sources. Because the effect of X-rays is the same as the case when reionization is driven only by QSOs, the reader can refer back to § 3.4.2. Although a QSO supplies an additional source of heating, its impacts on the thermal structure is very complex and nonlinear when the IGM is reionized by both galaxies and a QSO. We dissect the impact step by step.

First of all, as shown in the radial temperature profiles (Figure 3.20), a QSO increases the radius at which the gas is heated to $\sim 10^4$ K as expected. However, similar to the case for galaxies only models, galaxies+QSO models show a lower temperature in the inner parts of the H II region as well as similar overall temperature-density relation (Figure 3.19). This means that a QSO plays a small role in controlling the thermal state of the inner parts of the H II region where the pre-existing galaxies have already ionized the IGM before the QSO activity.

A QSO does, however, provide an additional heating at detectable level by eye. The top panels in Figure 3.19 show that galaxies+QSO models attain a higher temperature than galaxies only models at a fixed total ionizing emissivity. The gas heated by a QSO can generally reach $\log_{10} T/\mathrm{K} \gtrsim 4.2$; a large fraction of this high temperature gas is absent in galaxies only models. This gas appears to decrease (from left to right panels in three

[11]Note that because the photoionization of the residual neutral gas inside H II region can contribute to heating, the temperature toward the inner regions does not decrease monotonically.

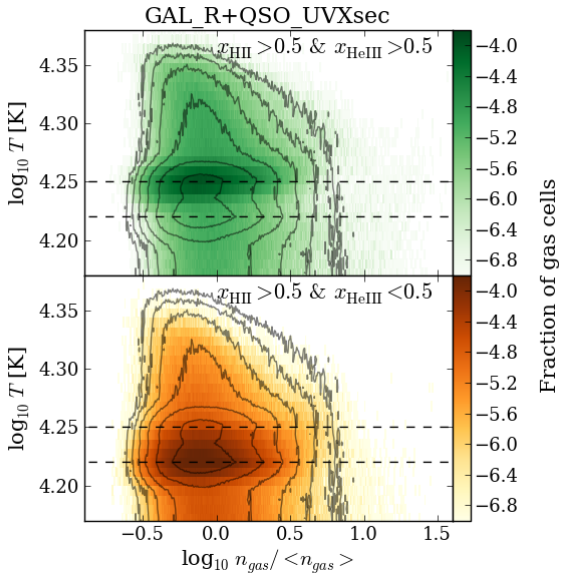

Figure 3.21: The zoom-in temperature-density diagram of the H II region in galaxies+QSO model GAL_R+QSO_UVXsec. The gas is separated according to the local H II and He III fractions: the gas with $x_{\rm H\,II} > 0.5$ and $x_{\rm He\,III} > 0.5$ (top) and the gas with $x_{\rm H\,II} > 0.5$ and $x_{\rm He\,III} < 0.5$ (bottom). As a guide, the black contours is the temperature-density diagram for the entire H II region (as in Figure 3.19. The horizontal dashed lines indicate the most frequent gas temperature in each region. The figure shows the bimodality in the temperature distribution.

galaxies+QSO models) as the contribution from galaxies increases (i.e. increasing total ionizing emissivity of galaxies).

To understand this, Figure 3.21 shows the zoom-in temperature-density diagram of the H II region in our reference galaxies+QSO model (GAL_R+QSO_UVXsec). The black contours indicate the relation in the entire H II region ($x_{\rm H\,II} > 0.5$) corresponding to the second rightmost panel in Figure 3.19. The H II region is further divided into two subregions; one inside the He III region ($x_{\rm H\,II} > 0.5$ and $x_{\rm He\,III} > 0.5$, green) and another outside the He III region ($x_{\rm H\,II} > 0.5$ and $x_{\rm He\,III} < 0.5$, yellow). The temperature distribution is bimodal, i.e. two population of gas at $\log_{10} T/{\rm K} \sim 4.25$ and $\log_{10} T/{\rm K} \sim 4.22$, which are seperated by the He III I-front inside the H II region. When both H II and He III I-fronts from a QSO exceed the pre-existing H II region of the surrounding galaxies (green plot in Figure 3.21), the gas experiences the most heating. On the other hand, when only the H II I-front from a QSO exceeds the pre-existing H II region, the He III I-front of a QSO is still smaller than the pre-existing H II region of galaxies (yellow plot in Figure 3.21). In this case, the gas newly heated up by H I photoionization of a QSO has not yet undergone QSO-driven He II photoionization. This lack of He II photoionization by a QSO results in a slightly lower temperature. This nonlinear heating mechanism, which depends on whether the He III I-front of a QSO is larger than the pre-existing H II I-front of the surrounding galaxies or not,

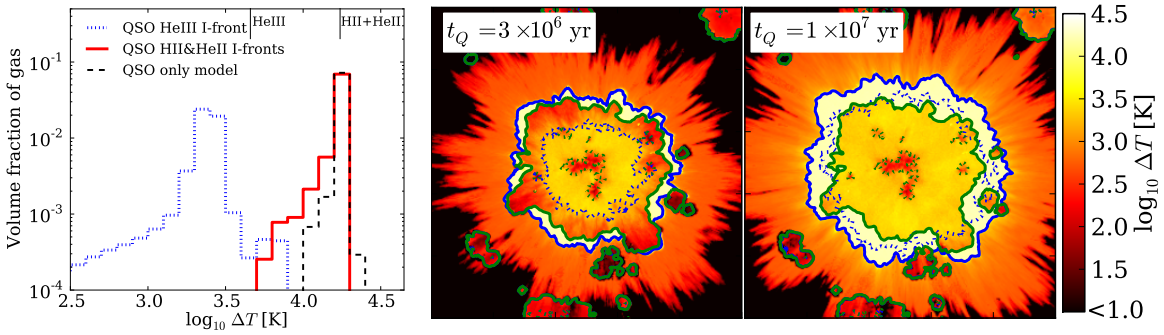

Figure 3.22: *Right*: The difference in the temperature between GAL_R+QSO_UVXsec model and GAL_R model, showing the additional heating by a QSO with two different QSO lifetimes: (left panel) $t_Q = 3 \times 10^6$ yr and (right panel) $t_Q = 1 \times 10^7$ yr. The blue solid (dotted) contour indicates the location of the QSO's H II (He III) I-front, and the green solid (dotted) contour indicates the location of the H II (He III) I-front already placed by pre-existing galaxies around a QSO. *Left*: The histogram showing the fraction of gas cells in simulations heated by ΔT by a QSO. The red solid (blue dotted) histogram shows the distributions of the temperature jump caused by the additional QSO heating after the H II (He III) I-front passed through. For a comparison, the black dashed histogram shows the distributions of temperature jump in QSO only model (QSO_UVXsec) across the H II I-front. The two vertical lines are the analytic estimates given in § 3.4.1. These figures demonstrate the nonlinear thermal impact of the QSO in a galaxy overdensity as only the region swept by QSO's I-front heated most (see text for detail).

explains why the QSO heating is ineffective for an increasing contribution from galaxies; larger the pre-existing H II region by galaxies is, lesser the region of gas is engulfed by both H II and He III I-fronts of a QSO.

It is worth elaborating on the collective nonlinear thermal effect when the IGM is reionized both by galaxies and QSOs. To examine the additional heating by a QSO in the galaxy overdensity, Figure 3.22 (right) shows differential maps of temperature between galaxies+QSO model (GAL_R+QSO_UVXsec) and galaxies only model (GAL_R) at two different QSO lifetimes. The solid blue and green contours indicate the location of H II I-fronts of a QSO and galaxies, respectively. Similarly the dotted blue and green contours correspond to the location of their He III I-fronts.

The structure of the additional QSO heating is clearly associated with the relative positions of the QSO H II/He III I-fronts and those of the pre-existing galaxies. The gas newly swept by the QSO H II I-front (between the blue solid and green solid contours) is heated up most (by $\Delta T \sim 17000$ K). The gas newly swept by the QSO He III I-front (between the blue dashed and green dashed contours) is heated up second most (by $\Delta T \sim 3000$ K). Note that at $t_Q = 3 \times 10^6$ yr the radius of the QSO's He III region has not yet exceeded the initial radius of the H II region produced by the pre-exisiting galaxies.

This means that the region between the galaxies' H II I-front (green solid contour) and the QSO's He III I-front (blue dashed contour) experiences little heating by a QSO (only by $\Delta T \sim 300$ K). This small heating is caused by the partial ionization of He II to He III by the X-ray photons. At $t_Q = 10^7$ yr the QSO He III region has grown large enough to heat up all the H II regions which are initially heated by the pre-existing galaxies. It is worth noting that when a QSO shines in a region reionized already by galaxies, it adds the heat mainly by the QSO-driven He II reionization inside the pre-existing H II region of galaxies, whereas the QSO-driven H I reionization can add a large amount of heat outside H II region of galaxies.

This nonlinear impact of the QSO heating can be understood in terms of the nature of photoionization heating and the spectral hardness of sources. First, in order for photoionization to heat up the gas, there must be neutral gas ahead of the I-front. This means that once the gas is ionized by galaxies, a QSO can contribute to heating little by photoionization of residual neutral gas. The amount of photo-heating is thus already set by the sources which first ionize the region (for example, by pre-existing galaxies). Thus, the QSO heating is effective only in the region that is newly ionized by the QSO. Second, the difference between galaxies and QSO models is the hardness of spectrum. The QSO I-fronts can more effectively dump thermal energy than galaxies (by about a factor of $\Delta T^{\rm QSO}/\Delta T^{\rm GAL} \approx (2+\alpha_{\rm GAL})/(2+\alpha_{\rm QSO}) \approx 1.4$). Thus, the gas *only* swept by the QSO's I-fronts can attain higher temperature than the gas first swept by the I-front of galaxies and QSO's I-fronts following after.

To make a more quantitative analysis, Figure 3.22 (left) shows the histograms of the fraction of gas heated by the passage of the QSO's H II (red solid) and He III (blue dotted) I-fronts. Because He I reionization occurs almost simultanously to H I reionization, the position of He II I-front is identical to that of H II I-front. For a comparison, the black dashed histogram shows the fraction of gas heated across the passage of the H II I-front in the QSO only model (QSO_UVXsec). The vertical solid lines show the analytic expectation of the temperature jumps caused by each I-fronts (§ 3.4.1). The QSO H II and He II I-fronts indeed increase the temperature by $\Delta T \sim 17000$ K. This is well captured by the analytic estimate of H I and He I photoionization heating. The fraction of gas heated in the galaxies+QSO model is similar to that in the QSO only model. This is not surprising. Because the luminosity of the QSO is identical in the two models, it can ionize and photoheat the same volume of the gas, although with different geometry. On the other hand, the QSO He III I-front increases the temperature by ~ 3000 K. Although this is somewhat lower than the simple analytic esimate, the both results from the analytic estimate and the simulations support the physical picture that the passage of the QSO He III I-front is a major source of heating inside the pre-existing H II region of galaxies.

In summary, the hydrogen and helium reionization driven by galaxies and QSOs nonlinearly impacts the thermal structure of the high-z QSO environment, which substantially differs from a simplier reionization model driven only by galaxies or QSOs. The relative extent to which the H II I-front of galaxies and the H II/He III I-fronts of QSO extend determines the thermal structure of the IGM. In general, higher temperature can be achieved if the gas is heated *only* with QSOs (harder spectrum); when more gas is ionized by galaxies,

the central QSO less efficiently heats up the IGM in the QSO environment.

3.5 Comparison with previous work

We compare our results with previous works and try to build a uniform picture to reconcile some of the differences with previous works. This (hopefully) provides a solution to some of the outstanding problems regarding the role of QSOs during reionization.

3.5.1 Can a QSO change the morphology of H II regions in a distinctive way?

Yes and no. The answer depends on (1) the relative amount of the total ionizing photons emitted by a QSO and galaxies since the onset of reionization and (2) the spectrum of a QSO. We settle this recursively raised issue of whether a QSO creates a totally distinguishable morphology (size and shape) of H II regions from that of galaxies.

The importance of the ratio of the total ionizing photons between galaxies, $N_{ion}^{\rm GAL}$, and a QSO, $\dot{N}_{ion}^{\rm QSO} t_Q$, is also noted by [103] (see our § 3.3.3 and [254]). As a QSO is likely born in a galaxy overdensity as predicted by previous analytical and numerical works ([254, 169, 493, 7]), the effect of the surrounding galaxies should be taken into account. Our results support this picture. While a QSO dominates the local ionizing photon emissivity during its activity, galaxies can participate to reionization for a longer period of time. Therefore, the morphology of H II regions is determined by the total amount of ionizing photons emitted by QSOs and galaxies 'since the onset of reionization'.

When the total ionizing photons from a QSO dominate over those of the surrounding galaxies ($\dot{N}_{ion}^{\rm QSO} t_Q > N_{ion}^{\rm GAL}$), the QSO is a main source of reionization in its local environment. This limit corresponds to the scenarios considered by [143, 444] and our models discussed in § 3.3.1. As galaxies are subdominant in these models, the QSO-dominated reionization produces a large spherical H II region distinctive from that of galaxies. The difference between these works and the results by [103] (and our galaxies+QSO models), which claim an important contribution from galaxies, boils down to the difference in the $\dot{N}_{ion}^{\rm QSO} t_Q$ to $N_{ion}^{\rm GAL}$ ratio. Because [143, 444] and our models in § 3.3.1 assume that both QSO and galaxies drive reionization for the same period of time, the contribution from galaxies is estimated to be much smaller than the cases of [103] and our galaxies+QSO models. This also means that at earlier times a QSO can more easily imprint a distinctive H II region than late times ([169]) because of a smaller cumulative contribution from galaxies.

When the total ionizing photons from a QSO are comparable to (or less than) those of surrounding galaxies [$\dot{N}_{ion}^{\rm QSO} t_Q \sim N_{ion}^{\rm GAL}$ (or $\dot{N}_{ion}^{\rm QSO} t_Q < N_{ion}^{\rm GAL}$)], it is difficult for a QSO to dominate the reionization of the local environment. This corresponds to the scenario considered by [103, 254] and our galaxies+QSO models discussed in § 3.3.3. In this case, a QSO does not change the morphology of the H II region drastically, but slightly enhances the degree of sphericality of the H II region with increasing ionizing luminosity of the QSO.

Because the partial ionization by X-rays from a QSO raises the H II fraction to only $\sim 10^{-2}$, the presence of QSO's X-ray photons alone cannot change the H II region drastically.

As the contribution from a QSO increases, the H II region becomes more spherical (a distinctive signature of a QSO). However, the breakdown of the sphericality occurs (1) when the QSO I-front merges with the initially unconnected H II regions of galaxies ([143]) and (2) when the QSO I-front hits the high density clumps of the gas, e.g. Lyman-limit systems, which cast shadow behind (see § 3.3.1 & 3.3.3 and [221]). Both factors imply that the degree of sphericality is influenced both by the QSO-galaxy clustering and QSO-absorber clustering ([221]) (incident rate of optically thick absorbers around a QSO).

The spectrum of a QSO plays a pivotal role in determining the morphology of H II regions. By studying QSO models with UV-obscured and unobscured spectra, previous works ([236] and also [498, 444]) concluded that the obscured QSO spectrum produces a thick I-front. Thus, as a QSO spectrum becomes more UV-obscured and harder, it can imprint a distinctive ionization structure, providing smoother and more extended I-front. On the other hand, a (unobscured) QSO emitting UV photons makes QSO's I-front looks more similar to galaxies. Because the penetrating power of typical photons from a UV-obscured spectrum is high, the QSO's H II region is more spherical. Our result is consistent with this picture, but adds an another ingredient. Although for a single LOS, a harder and UV-obscured spectrum thickens the I-front, in many LOSs there is substantial fluctuation of I-front position due to the density fluctuation of the IGM and a possible merger with the H II regions of galaxies. Therefore, in order for the thickness of the I-front to be used as an indicator of sources, the thickening of the I-front must exceed the width due to the LOS fluctuation of I-fronts. In addition, soft X-ray photons from a QSO must dominate the growth of I-fronts over the contribution from the UV photons of pre-existing galaxies in order for an obscured QSO to imprint a distinctive morphology of the H II region.

The above discussion concludes the answer to the question: Yes, a QSO does imprint a distinctive spherical hydrogen ionization structure for a reionization scenario when a QSO has a hard UV-obscured spectrum and/or a QSO dominates the total ionizing photon budget in its local environment over that of galaxies integrated since the onset of reionization. No, when a QSO has an unobscured spectrum and the total ionizing photon budget of surrounding galaxies is more than or comparable to that of the central QSO. We argue that given a possible parameter space of high-z galaxy and QSO properties, it would be premature to associate a large spherical H II region with the presence of a QSO inside. The morphology of the H II regions is a complex product depending on the QSO's spectral type, surrounding galaxies, and the density fluctuations of the IGM in the environment.

3.6 Observational Implications

A natural question in the EoR science is how we can learn about the role of galaxies and QSOs in driving reionization. In this chapter we have focused on the theoretical aspects of reionization models driven by galaxies and QSOs. We briefly discuss the observational implications of our theoretical insights.

A direct future observational strategy for studying the sources of reionization is to perform a galaxy survey in a region of the sky imaged by 21cm tomography. It has been well-known that 21cm image traces directly the morphology of H II regions during reionization (e.g. [159, 312, 298]). Although it is true that the structure of a QSO H II region, which appears as a coldspot in 21cm image, contains information about the ionizing sources ([487, 444, 236, 103]), as emphasized in this work, the morphology of H II regions in high-z QSO environments is complicated by surrounding galaxies and spectral shape of QSOs. This implies that there will still remain a large degenerarcy when constraining the sources using 21cm data alone. Galaxy surveys (via dropout or narrow-band technique) in the direction of coldspots in 21cm tomography provide an obvious complementary information. This strategy can be realized, for example, in ELAIS-N1 field targeted by LOFAR ([211]) and Subaru/Hyper-SuprimeCam Survey[12]. As 21cm experiments require the deepest ever radio sky observation, one may employ a deep radio extragalactic survey ([29, 462]) which comes together with the data aimed for 21cm experiment. Alternatively, 21cm images can be formed in the direction of high redshift QSOs after they are found by galaxy surveys. Ongoing wide field surveys such as Dark Energy Survey and Subaru/Hyper-SuprimeCam Survey are expected to provide $\sim$ 10 new QSOs at $z \sim 7$ ([367, 272]). This approach allows us to directly infer the expected ionizing photon counts emitted from galaxies using surveys, together with the their impact on the morphology (e.g. size) of their H II regions (coldspots) using 21cm tomography. Therefore, this directly probes whether the estimated ionizing photon counts of galaxies and QSOs are sufficient to create the observed size of the H II region harbouring the sources. For example, our analytic model presented in § 3.3.3 can be used to help the interpretation of such a galaxy survey plus 21cm tomography observations. This strategy is arguably the most direct test of the sources of reionization, i.e. how galaxies and/or QSOs reionized the IGM.

Another strategy is to perform a galaxy survey in the regions around high redshift QSOs. The properties of galaxies and QSOs during reionization and the physical state of the IGM can be directly studied by the photometric/spectroscopic data from the survey and Lyα absorption in the QSO spectra, respectively. An advantage of this strategy is that it can be realized already using current observational facilities, although the interpretation is more complex than the above galaxy survey plus 21cm tomography strategy. Observing the ionizing luminosity and the escape fractions of high-redshift galaxies and QSOs alone does not immediately imply that they are 'drivers' of reionization. They may be shining in a region already reionized by other sources. As demonstrated in our work, the impacts of galaxies and QSOs (as traced by the galaxy survey) on the ionization and thermal states of the IGM (as traced by e.g. QSO near-zone and Lyα absorption line widths) are important complementary information about how they have driven reionization. For example, [80] used the temperature measurement in the local environment of $z \sim 6$ QSOs by [41] to derive constraints on the ionizing sources. In this sense, the data obtained by this strategy (galaxy survey in QSO fields) is perhaps best exploited once the interpretation is aided by a RT simulation suite.

[12] http://www.naoj.org/Projects/HSC/surveyplan.html

3.7 Conclusions

We have presented detailed analysis of hydrogen and helium reionization scenarios in high-z QSO environments including both galaxies and a central QSO using a suite of cosmological multi-frequency radiative transfer simulations. This allows us to understand the array of physical mechanisms that shapes the ionization and thermal states of the IGM. We summarize the role of QSOs in galaxy overdensities in driving reionization and thermal state of the IGM.

We find that, when reionization is driven by galaxies and QSOs, the morphology of H II regions in high-z QSO environments is complicated by the presence of galaxies around a central QSO. While a single QSO can dominate the ionizing power in its vicinity during the active phase, the surrounding galaxies can reionize a substantial region of the QSO environment before the QSO activity. Thus, if the total integrated ionizing photon emitted from galaxies since the onset of reionization exceeds or is comparable to that of QSOs, the distinction in the reionization morphology between galaxies only and galaxies+QSO models become vague. On the other hand, if QSOs dominate the total integrated budget around the environments, QSOs can produce a large distinctive morphology of H II regions.

The unique impact of a high-redshift QSO comes from its hard spectrum. As a result, QSOs may drive the intial phase of He II reionization at redshift as high as the epoch of H I reionization. Therefore, the morphology of He II and He III regions is a distinctive and unique signature of QSOs compared to that of galaxies.

In addition, X-ray photons from QSOs produce extended partially ionized tails of hydrogen and helium around the central H II and He II/He III regions. Secondary ionization enhances the contribution of X-ray photons to reionization at the partially ionized tail. On the other hand, the morhopology of fully ionized regions bounded by sharp I-fronts is insensitive to the X-ray photons. The I-front thickness is not affected significantly by X-ray physics because the UV photons from the surrounding galaxies or the central QSO itself are already sufficient to produce sharp I-fronts.

Different intrinsic properties of QSOs influence differently the ionization states of hydrogen and helium in the IGM. Firstly, the spectral shape of QSOs plays an important role. UV-obscured QSOs can broaden the I-fronts significantly. The noticeable change in the ionization morphology occurs when the soft X-ray photons from a QSO dominate the growth of cosmological I-fronts over the UV photons from galaxies. In other words, the net composite spectral shape of the collective radiation field from galaxies and QSOs driving reionization must be X-ray dominant in order for obscured QSOs to produce distinctive H II regions in their local environments.

QSO duty cycle also plays an important role. In order for QSOs to contribute reionization of hydrogen and helium, QSO activities must be continuously fuelled so that a next QSO activity occurs before the ionizing photons from the previous activity recombine. QSO-driven H II regions can survive when $z > 6$ QSOs are fuelled by frequent minor mergers. On the other hand, if the QSOs are fuelled by less frequent major mergers, they must repeatedly make the H II regions from scratch at every activity; in this case, the contribution of QSOs to reionization is less significant. The continous fuelling of QSOs must

occur even more vigorously for helium reionization than hydrogen to sustain the onset of He II reionization at $z > 6$.

The thermal state of the IGM is then strongly affected by the photoionization heating of both hydrogen and helium associated with the reionization process driven by galaxies and QSOs. A central QSO and the surrounding galaxies in high-z QSO environments non-linearly impacts the thermal state of the post-reionized IGM. The thermal structure depends on the 'relative' contributions from galaxies and QSOs to the total radiation field driving reionization. A larger amount of total ionizing photons does not necessarily increase the gas temperature. The highest temperature can be attained in the region ionized only by QSOs. More regions the galaxies ionize, less efficient the QSO heating becomes.

The X-ray photons from a central QSO heat the gas ahead of I-fronts to $T \sim 10^3$ K at the extended tail of partial ionization. The net X-ray heating is slightly reduced when secondary ionization is included because it adds another channel to deposit the energy into ionization. The X-ray pre-heating is a clean signature of QSOs.

In summary, based on the above findings, we argue *the concerted scenario of hydrogen and helium reionization driven by both galaxies and QSOs in high-z QSO environments*. They collectively and non-linearly shapes the physical state of the IGM. This picture emphasizes the importance of treating hydrogen and helium reionization on an equal footing to understand the role of galaxies and QSO during reionization. To correctly model this scenario, multi-frequency treatment of radiative transfer simulations with both hydrogen and helium is as important as acheiving a high spatial resolution in radiation hydrodynamical simulations. We have demonstrated that it is now possible to produce a large suite of multi-frequency RT simulations. This provides a highly useful theoretical resource to understand the physical mechanisms responsible for shaping the ionization and thermal structures of the IGM. This will facilitate the use of RT simulations in interpreting QSO spectra and upcoming 21cm observations to better understand the role of galaxies and QSOs in driving reionization.

Chapter 4

Radio signals from the high redshift QSO environments

In Chapter 4 I present the H I 21cm and ^{3}He II 3.46cm signals from the intergalactic hydrogen and helium around the high redshift QSOs. These radio signals allow us to test the detail physics of the reionization process driven by galaxies and/or QSOs.

4.1 Backgrounds

21cm cosmology, i.e. looking for a radio signal from the hyperfine transition of neutral hydrogen in the intergalactic gas, promises unprecedented knowledge about Cosmic Dawn and the EoR. Many observational programmes such as LOFAR, MWA, PAPER, and GMRT, are underway to detect this 21cm signal from the high redshift IGM. Future radio interferometers such as the Square Kilometer Array (SKA) and HERA are designed to study the high-redshift universe by means of low frequency radio astronomy.

Motivated by this exciting possibility, a hyperfine 3.46cm (8.7 GHz) transition of singly ionized isotope of helium (^{3}He II) is also proposed as a probe of intergalactic helium ([159, 287, 11]). Although the primordial abundance of ^{3}He relative to hydrogen produced by Big Bang Nucleosynthesis is $\sim 10^{-5}$ ([429]), the high (by a factor of 680) spontaneous transition rate of ^{3}He II means that the hyperfine 3.46cm emission can be strong. In addition, because the sky is quieter for this frequency range, the sky noise of radio interferometric observations may be less problematic than that for 21cm signal.

The observability of H I 21cm and ^{3}He II 3.46cm signals using the low and mid frequency radio interferometers permits us to study the reionization process of hydrogen and helium driven by the early population of galaxies and black holes in a great detail.

4.2 Cosmological radio signals of the IGM

4.2.1 H I 21 cm signal

The hydrogen hyperfine ground state due to the spin flip transition corresponds to the 21cm line with $\nu_{21} = 1420.4$ MHz, whose emission or absorption at redshift z is observed at a frequency, $1420.4(1+z)^{-1}$MHz. The differential brightness temperature of 21cm line against the CMB is given by (e.g. [146, 147, 262])

$$\delta T_{b,\mathrm{H\,I}} = T_0(z)(1+\delta_b)x_{\mathrm{HI}}\left(1-\frac{T_{\mathrm{CMB}}(z)}{T_S}\right)\left(1+\frac{1}{H(z)}\frac{dv_\parallel}{dr_\parallel}\right)^{-1}, \tag{4.1}$$

where $T_0(z) = \frac{3h_p c^3 A_{21}}{32\pi k_{\mathrm{B}}\nu_{21}^2}\frac{(1+z)^2\bar{n}_{\mathrm{H}}(0)}{H(z)} \approx 28\ \mathrm{mK}\left(\frac{\Omega_b h^2}{0.023}\right)\sqrt{\frac{1+z}{10}\frac{0.15}{\Omega_m h^2}}$ depends on physical constants, cosmological parameters, and the mean number density of hydrogen. $\bar{n}_{\mathrm{H}}(0)$ is the comoving mean number density of hydrogen in the universe, $A_{21} = 2.85\times10^{-15}s^{-1}$ is the Einstein A-coefficient of the 21cm transition, and $H(z)$ is the Hubble parameter, and the other physical constants have usual meaning. The term involving the line-of-sight peculiar velocity gradient $dv_\parallel/dr_\parallel$ is the effect of redshift space distortion. $T_{\mathrm{CMB}}(z) = 2.73(1+z)$ K is the CMB temperature at redshift z.

The spin temperature T_S encodes the detailed physics of 21cm line transition, which is given by (e.g. [146, 159])

$$T_S^{-1} = \frac{T_{\mathrm{CMB}}^{-1} + x_\alpha T_\alpha^{-1} + x_c T_K^{-1}}{1+x_\alpha+x_c}, \tag{4.2}$$

where the collisional coupling coefficient, x_c, and the Lyα coupling coefficient, x_α, are determined by the collisional process of hydrogen atoms and radiative transition due to Lyα scattering, called the Wouthuysen-Field mechanism ([485, 146]). During the EoR, the collisional coupling is only effective at high density clumps. The Lyα coupling by the Wouthuysen-Field effect plays a dominant role. The Lyα background can be built up based on a reasonable assumption about the population of early star forming galaxies (e.g. [262, 159, 354]). It couples the spin temperature to the color temperature $T_S \to T_\alpha$, and the color temperature T_α to the kinetic temperature of the gas $T_\alpha = T_K$. We assume this coupling in this chapter, which simplifies the differential brightness temperature as,

$$\delta T_{b,\mathrm{H\,I}} \approx T_0(z)(1+\delta_b)x_{\mathrm{HI}}\left(1-\frac{T_{\mathrm{CMB}}(z)}{T_K}\right). \tag{4.3}$$

We have also ignored the effect of the redshift space distortion by the peculiar velocity.

4.2.2 ^{3}He II 3.46 cm signal

The underlying physics for ^{3}He II 3.46 cm signal is analogous to the H I 21 cm signal. Following the same line of thought, the differential brightness temperature of ^{3}He II 3.46

cm against the CMB is given by (e.g. [159]),

$$\delta T_{b,^3\mathrm{He}} \approx 1.76 x_{\mathrm{He\,II}}(1+\delta_b)\left(\frac{y_{^3\mathrm{He}}}{10^{-5}}\right)\left(1-\frac{T_{\mathrm{CMB}}(z)}{T_{S,^3\mathrm{He}}}\right)\left(1+\frac{1}{H(z)}\frac{dv_\parallel}{dr_\parallel}\right)^{-1}\left(\frac{1+z}{11}\right)^{1/2}\mu\mathrm{K} \tag{4.4}$$

where $y_{^3\mathrm{He}} \approx 1.0\times10^{-5}$ is the primordial abundance ratio (by number) of helium isotope-3 relative to hydrogen atoms by big bang neucleosynthesis ([429]). Helium isotope-4 (^{4}He) has $y_{^4\mathrm{He}} = 0.083$, which dominates the total abundance of helium.

The spin temperature of ^{3}He II is determined by the collisional process and He II Lyα radiative transfer via the He II analog of the Wouthuysen-Field effect ([287, 11, 434]),

$$T^{-1}_{S,^3\mathrm{He}} = \frac{T^{-1}_{\mathrm{CMB}} + x_{\alpha,^3\mathrm{He}}T^{-1}_{\alpha,^3\mathrm{He}} + x_{c,^3\mathrm{He}}T^{-1}_{K}}{1 + x_{\alpha,^3\mathrm{He}} + x_{c,^3\mathrm{He}}}, \tag{4.5}$$

where $x_{c,^3\mathrm{He}}$ and $x_{\alpha,^3\mathrm{He}}$ are the collisional and He II Lyα coupling coefficients, respectively. $T_{\alpha,^3\mathrm{He}}$ is the color temperature of the radiation field near the ^{3}He II Lyα line. Similar to the H I 21cm line, the collisional coupling is only efficient at high density clumps $\Delta_b \gtrsim 100$ ([287, 11]).

The ^{3}He II 3.46 cm line can also be pumped by the He II analog of the Wouthuysen-Field mechanism by scattering off of UV photons at He II Lyα line (40.8 eV). However, the efficiency of the ^{3}He II Wouthuysen-Field mechanism is less clear than the H I case. Because of the presence of more abundant ^{4}He II by a factor of $y_{^4\mathrm{He}}/y_{^3\mathrm{He}} \approx 7.5\times10^3$, most of He II Ly$\alpha$ scatterings is caused by ^{4}He II. This led previous studies ([79, 287]) to conclude that the Wouthuysen-Field mechanism is inefficient for ^{3}He II. However, this ignores the effect of multiple scatterings of He II Lyα photons, whose He II Gunn-Peterson optical depth is high $\tau^{\mathrm{He\,II}}_{\mathrm{GP}} = (y_{^4\mathrm{He}}/4)\tau^{\mathrm{H\,I}}_{\mathrm{GP}} \approx 1.6\times10^4[(1+z)/11]^{3/2}$. Thus, $\tau^{\mathrm{He\,II}}_{\mathrm{GP}} \approx 1.6\times10^4$ scatterings of ^{4}He II can effectively compensate the low abundance of ^{3}He II as $\tau^{\mathrm{He\,II}}_{\mathrm{GP}}(y_{^3\mathrm{He}}/y_{^4\mathrm{He}}) \approx 2$. This can boost the previous estimate of the He II Wouthuysen-Field effect by a factor of 2.

In addition, the previous estimates ([79, 434]) have considered the Wouthuysen-Field coupling strength by a homogeneous UV background. While this is relevant for a typical region of the diffuse IGM, we consider the IGM around QSOs. In the QSO environment, the high energy UV photons blueward of He II Lyα line ($>$ 40.8eV) can be provided by the central QSOs. This provides a large reservoir of He II Lyα radiation than the typical region ([79, 434]).

Therefore, in QSO environments, the Wouthuysen-Field coupling strength may not be as low as previously estimated, although a detailed modelling by solving the He II Lyα radiative transfer in an expanding universe is necessary to draw a quantitative conclusion.[1] In this chapter we take an optimistic assumption that the Wouthuysen-Field coupling is

[1]The color temperature of He II Lyα radiation field may not be fully coupled to the spin temperature ([109]). In practice, a full coupling is not necessary to observe the differential brightness temperature in emission as long as the coupling rises the color temperature sufficiently above the CMB temperature. I also note that the presence of O III removes the He II Lyα photons by the Bowen fluorescence ([109]), although I expect that this is less problematic for the diffuse IGM with low metal enrichment.

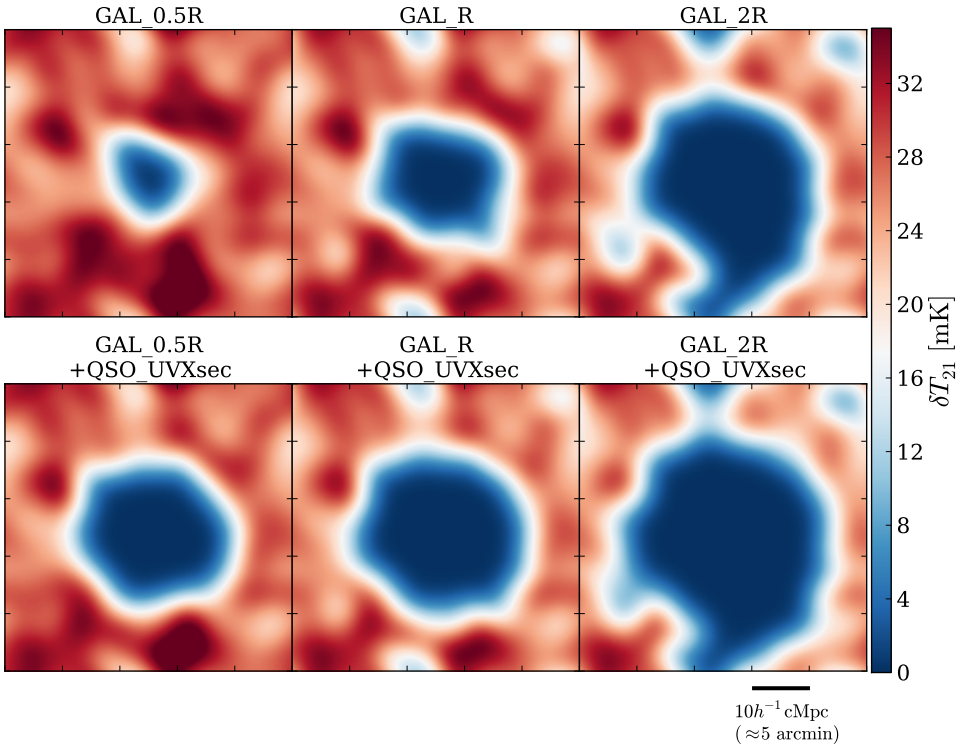

Figure 4.1: The 21cm signal around the galaxy overdensities (top panels: GAL_0.5R, GAL_R, GAL_2R) and $z = 10$ QSO environments with surrounding galaxies (bottom panels: GAL_0.5R+QSO_UVXsec, GAL_R+QSO_UVXsec, GAL_2R+QSO_UVXsec). The image is smoothed with angular resolution of 3.19 arcmin.

efficient so that the spin temperature is coupled to the color temperature, and the color temperature is coupled to the kinetic temperature,

$$\delta T_{b,^3\mathrm{He}} \approx 1.76 x_{\mathrm{He\,II}}(1+\delta_b)\left(\frac{y_{^3\mathrm{He}}}{10^{-5}}\right)\left(1-\frac{T_{\mathrm{CMB}}(z)}{T_K}\right)\left(\frac{1+z}{11}\right)^{1/2}\mu\mathrm{K}, \tag{4.6}$$

where we have also ignored the effect of the peculiar velocity. I note that while this is an optimistic upper estimate of the ^{3}He II 3.46 cm signal, it allows us to explore the ^{3}He II 3.46 cm signal using the results of the radiative transfer simulations for the first time.

4.3 H I 21 cm signal around high redshift QSOs

In this section, I discuss the properties of the H I 21cm signal around high redshift QSOs using the suite of radiative transfer simulations described in Chapter 3.

4.3.1 The full coupling limit

I first explore the 21cm signal in the full coupling limit. This assumes that the kinetic temperature of the gas is always above the CMB temperature $T_K \gg T_{\rm CMB}$. In this full coupling limit, the 21cm signal traces the neutral hydrogen density $\delta T_{b,\rm H\,I} \propto x_{\rm HI}(1+\delta_b)$. The full coupling assumption is conventionally used for studying the 21cm signal from the EoR, although it assumes the presence of efficient heating sources such as X-ray binaries or any other before the onset of reionization.

Figure 4.1 shows the simulated 21cm signals from the radiative transfer simulation suite. The top panels shows the results for the galaxies only models with the three different total ionizing photon emissivities (GAL_0.5R, GAL_R, GAL_2R). The bottom panels shows the results for the QSO environments at $z = 10$, i.e. the galaxies+QSO models (GAL_0.5R+QSO_UVXsec, GAL_R+QSO_UVXsec, GAL_2R+QSO_UVXsec). The images are smoothed at the angular resolution of 3.19 arcmin, which corresponds to the angular resolution achieved by using maximum 2 km baseline arrays of a radio interferometer such as LOFAR and SKA.

The figure shows that the central H II region appears as a coldspot in the 21cm image, i.e. deficit of the brightness temperature. Regardless to the type of sources, galaxies and/or QSOs, responsible for driving reionization, the 21cm coldspot simply appears as a spherical blob. A central QSO produces a tail of partially ionized region ahead of the H I I-front as described in Chapter 3. However, because the partially ionized IGM has only an ionized fraction of $x_{\rm H\,II} \sim 10^{-3}$, the difference in the 21cm signal (neutral hydrogen structure) from the galaxies only model is marginal. At the level of precision accessible by ongoing and planned radio interferometers, the 21cm signal from galaxy overdensities and QSO environments appears indistinguishable.

4.3.2 The kinetic coupling limit

The full coupling limit takes a reasonable, but strong, assumption about the existence of heating sources in the pre-reionization era. In our current knowledge both theoretical and observational, this involves a large extrapolation of the low redshift properties of galaxies to high redshift $z > 6$. I now consider a scenario dropping the assumption of the full coupling limit. This is self-consistent with the physics assumed in the radiative transfer simulation suite, where the only available source of X-ray heating is the central QSO. We refer to this scenario as the kinetic coupling limit $\delta T_{b,\rm H\,I} \propto x_{\rm HI}(1+\delta_b)(1-T_{\rm CMB}/T_K)$.

Figure 4.2 shows the simulated 21cm signal around a $z = 10$ QSO environment in the kinetic coupling limit. The structure of the 21cm signal is clearly different from the full coupling limit. The ring-shaped 21cm signal around a QSO environment is evident. This is caused by the X-ray heating of the neutral (partially ionized to $x_{\rm H\,II} \sim 10^{-3}$) hydrogen. In the absence of other efficient heating sources other than the X-rays from the central QSO, the gas around the H II region hosting a QSO is heated to $\sim 10^{2-3}$ K and gradually fades away at outer radii. The 21cm emission gradually decreases as the temperature decreases. Eventually the 21cm appears in absorption as no heating source is present ([78, 74, 444]).

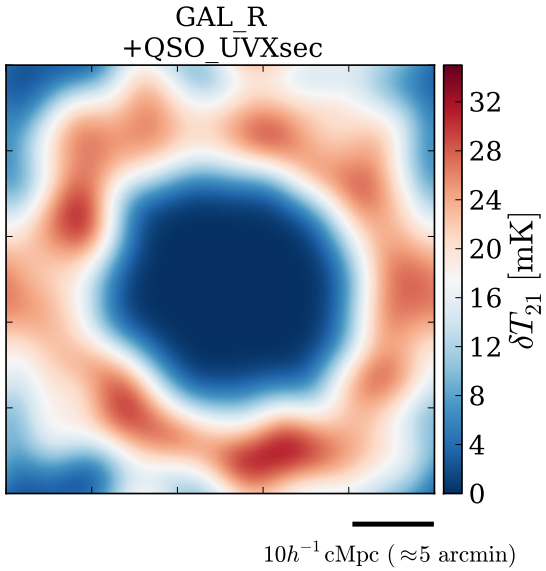

Figure 4.2: The simulated 21cm signal in the kinetic coupling limit, where only the source of pre-heating is the X-ray from the central QSO. The 21cm signal around $z = 10$ QSO environments with surrounding galaxies (GAL_R+QSO_UVXsec).

The cold IGM, i.e. 21cm in absorption, may be present during the reionization epoch ([145, 10]). The distinguishability of the 21cm signature (structure of 21cm coldspot) thus depends on whether there are other efficient sources of heating apart from QSOs. If a central QSO is only a source of pre-heating of the neutral IGM, it can imprint a unique signature on the structure of the 21cm signal. On the other hand, if other heating sources such as X-ray binaries are abundant, the 21cm signals from galaxy overdensities and QSO environments becomes indistinguishable.

4.4 ^{3}He II 3.46 cm signal around high redshift QSOs

The ^{3}He II 3.46cm signal from the intergalactic helium complements H I 21cm signal from the intergalactic hydrogen. Firstly, because the ^{3}He II 3.46cm signal probes He II reionization, the signal can isolate the impact of QSOs during reionization. In addition, because He I reionization occurs simultaneously with H I reionization also for a galaxy-driven reionization scenario, the ^{3}He II signal allows us to study the density fluctuations inside H II regions.

As noted in § 4.2.2, the physics of the ^{3}He II spin temperature is more complex than that of the H I 21cm line. I take an optimistic assumption of the kinetic coupling limit $\delta T_{b,^3\mathrm{He}} \propto x_{\mathrm{He\,II}}(1+\delta_b)(1-T_{\mathrm{CMB}}/T_K)$ for both galaxies only and galaxies+QSO models. Note that because the He II region is enclosed by H I I-fronts, the intergalactic helium is heated up to high ($> 10^4$K) due to the photoionization heating across H I I-fronts. Thus the kinetic coupling limit corresponds to the full coupling limit, $\delta T_{b,^3\mathrm{He}} \propto x_{\mathrm{He\,II}}(1+\delta_b)$,

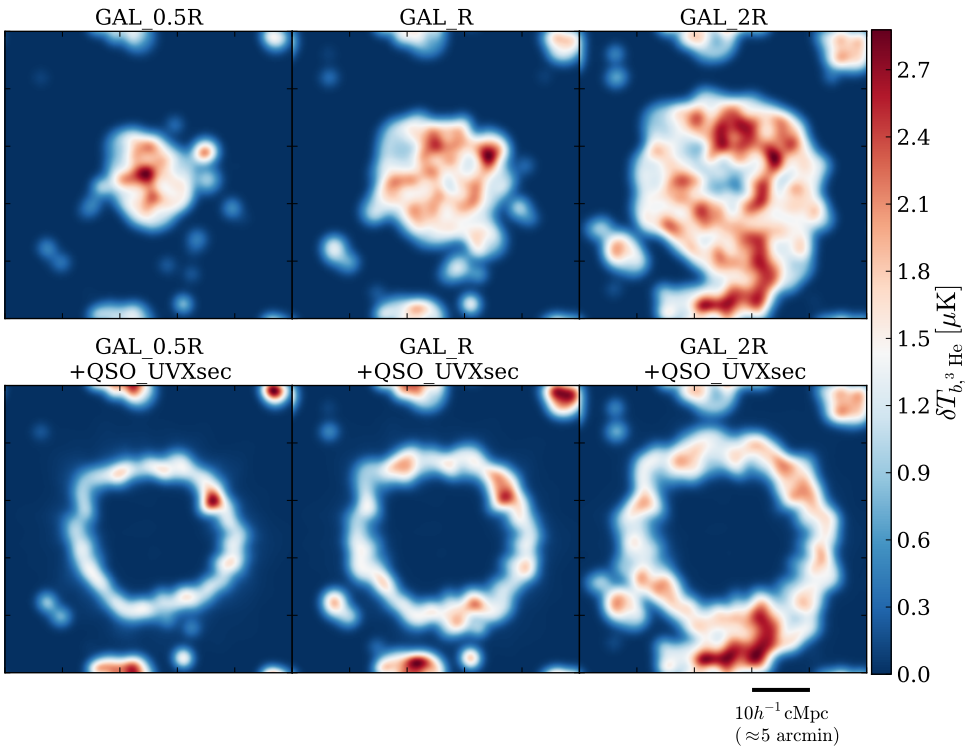

Figure 4.3: The ^{3}He II 3.46cm signals around the galaxy overdensities (top panels: GAL_0.5R, GAL_R, GAL_2R) and $z = 10$ QSO environments with surrounding galaxies (bottom panels: GAL_0.5R+QSO_UVXsec, GAL_R+QSO_UVXsec, GAL_2R+QSO_UVXsec). The image is smoothed with angular resolution of 1.31 arcmin.

once it is achieved.

Figure 4.3 shows the simulated ^{3}He II 3.46 cm signals from galaxy overdensities (top panels) and QSO environments (bottom panels). The three columns correspond to the different total ionizing photon emissivities from galaxies. The images of the 3.46 cm signal assume the angular resolution of the maximum 1 km array distribution, which is easily achievable by SKA1-mid.

The figure shows that the ^{3}He II 3.46cm signal in the kinetic coupling limit shows a clear difference between the galaxy overdensities and QSO environments. The ^{3}He II 3.46 cm signal from the galaxy overdensities traces the structure inside H II regions (He II regions). On the other hand, because the He II reionization is driven by QSOs immediately after He I reionization, the QSO environments show the He III region, appearing as a deficit region of the 3.46 cm signal. The region of ^{3}He II 3.46 cm signal in emission is smaller for increasing contribution of QSOs to reionization (right to left panels). The earlier reionization by the pre-exiting galaxies around a QSO produces a large He II region simultaneously to the H II region. As a central QSO ionizes He II to He III inside the galaxy-driven He II region, the shell of He II region is larger for larger contribution from the surrounding galaxies.

In principle, the inner density structure of the H II region is more easily accessible by ^{3}He II 3.46cm signal because of the higher frequency. The angular resolution of a radio interferometer scales with $\propto \lambda/b_{\rm max}$ where λ is the observed wavelength and $b_{\rm max}$ is the maximum baseline length. Thus the angular resolution can be higher for 3.46 cm signal by a factor of 6 than 21 cm signal at a given maximum baseline length.

4.5 Discussions

I have discussed the properties of the H I 21cm and ^{3}He II 3.46cm signal from the intergalactic hydrogen and helium from high redshift QSO environments. The theoretical study of these hyperfine transition signals prepares us for observational programmes of ongoing and upcoming radio interferometers such as LOFAR and SKA. A motivation behind this analysis is to understand what we can learn from the targeted study of the individual 21cm signal around a QSO.

The 21cm signal shows little difference whether the H I reionization is driven by galaxies or QSOs in the full coupling limit. In both cases, the 21cm signal appears as a large spherical coldspot in the 21cm image. This is partially caused by the low angular resolution of planned radio interferometers. However, the similarly in 21cm signal has a fundamental physical origin. The distinguishability of the 21m signal depends on the presence of heating sources apart from the central QSOs. While the targeted study of the 21cm signal can narrow down the broad range of the properties of sources, it is difficult to fully constrain the source properties by assessing the structure of the 21cm signal alone. It will be best studied by follow-up infrared observations by JWST, which directly search galaxies or QSOs inside H II regions.

On the other hand, in principle, the ^{3}He II 3.46cm signal allows us to disentangle the source properties by looking at their impact on helium reionization. The signal is highly

complementary to the H I 21cm signal. The current obstacle for the observability of ^{3}He II 3.46cm signal originates from the subtle physics of the spin temperature; in particular, the efficiency of the ^{3}He II Wouthuysen-Field mechanism.

For the high redshift ^{3}He II 3.46cm signal at $z > 6$, the lower redshift ($z < 3$) 21cm emitting galaxies act as interlopers. However, the signals can be separated. The available template of the 3.46cm signal and its distinct shell-shaped signal can be used for the separation. The interloper 21cm emitting galaxies can also be directly removed by searching for the optical/infrared counterparts by JWST. Because the target field of view of 21cm coldspot around a known high redshift QSO is relatively small, the direct search may be performed with a reasonable integration time. Finally, while the even optimistic estimate of the ^{3}He II 3.46cm signal is μK level, the experiment can be performed as a piggy-bag project of low redshift 21cm intensity mapping aiming to study cosmology by the baryonic acoustic oscillation (BAO).

The synergy between infrared facilities such as JWST and radio facilities such as SKA1-low and -mid opens up an exciting observing strategy. Galaxy surveys in the foreground of high redshift QSOs at $z > 6$ will be highly valuable. Such high redshift QSO fields permit us to directly probe galaxies and QSOs using JWST as well as the intergalactic hydrogen and helium using the upcoming radio interferometry, SKA1-low/mid.

Chapter 5

Conclusions and Prospects

The study of the Epoch of Reionization (EoR) and high redshift galaxies in the first billion years in cosmic history has now been undergoing an era of tremendous progress. The era of first galaxy formation and reionization is a landmark cosmic event when the emergence of the first astronomical objects has transformed the global environment of the universe. Understanding the EoR is considered to be one of the most important frontiers in extragalactic astronomy and cosmology in the 21st century. It is indeed a key science driver for many future telescopes such as Square Kilometer Array, James Webb Space Telescope, European Extremely Large Telescope, Thirty Meter Telescope, and Giant Magellan Telescope.

We can now start understanding the truly uncharted territory of the high redshift universe using a combination of observational data and numerical simulations/theoretical modelling. The observational knowledge on reionization can be pushed beyond the two traditional probes (Gunn-Peterson trough in QSO spectra and the CMB) using Lyα emitting galaxies and 21cm cosmology. Ultra deep observations using space- and ground-based telescopes have routinely been revealing galaxies at $z > 6$, reaching a redshift as high as $z \sim 12$. An iconic high-redshift galaxy survey – Hubble Ultra Deep Field – is a notable example. Furthermore, high-redshift galaxy surveys can now be used to constrain the nature of early galaxies and the reionization process. For instance, surveys of Lyα emitting galaxies now occupy an important place as a reionization probe and to understand the high-redshift galaxy formation. In addition, absorption line studies of QSO and gamma-ray burst spectra have been providing new insights into the physical state of the IGM near and during the EoR. In observations of low frequency radio sky, the community of 21cm cosmology – a direct probe of reionization – has been progressing toward the first detection of the 21cm signal. Ever improving upper limits on the 21cm power spectrum have been reported. It is our challenge to put all these observations together into a coherent picture of the Epoch of Reionization and high-redshift galaxy formation since Cosmic Dawn.

Theoretical progress has now made possible to realistically model the reionization process. Cosmological radiation hydrodynamical simulations as well as multi-frequency radiative transfer simulations can be performed to theoretically investigate the reionization era. This allows us to understand the physical mechanisms driving reionization and the

properties of galaxies and QSOs during the epoch. A set of simulations and models are also crucial to interpret the complex data of galaxy surveys and 21cm experiments to correctly constrain the physics of the EoR.

Chapter 1 is devoted to set the foundation to explore the high-redshift universe; in particular, the Epoch of Reionization. The aspects of extragalactic astrophysics relevant for understanding the reionization and high-redshift galaxy formation are summarized. I discussed the current status of our knowledge and observational constraints on reionization, the high-redshift galaxy population, and the intergalactic medium.

In Chapter 2, I have used Lyα emitting galaxies as a probe of reionization. I have considered how and what can be learned about reionization – the reionization history and morphology – using surveys of Lyα emitting galaxies. I created a set of reionization models using radiative transfer simulations post-processing a cosmological hydrodynamical simulation. The model spans a wide range of possible ionization topologies of the IGM, namely various combinations of large-scale H II regions and small-scale absorbers such as Lyman limit systems and damped Lyα systems. I emphasize the importance of analysing the data using a reionization framework which contains small-scale absorbers within large-scale H II regions. This is a crucial step to correctly constrain the global neutral fraction of the IGM (reionization history) in an unbiased way.

By performing a careful interpretation of the data using simulations, I suggests that the current data likely favours a late and rapid reionization history, albeit with large both theoretical and observational uncertainties. This constraint on the reionization history is degenerate with the assumptions about the intrinsic properties of Lyα emitting galaxies and the large-scale (distribution of H II regions) and small-scale (e.g. Lyman limit systems and damped Lyα systems) ionization structure of the IGM. This means that both the history and topology of reionization must be probed simultaneously to derive a better constraint. The physical origin of the degeneracy is due to the clumpy gas distribution around galaxies. The clumpy gas acts as an additional intervening medium suppressing the Lyα flux from galaxies, which must be distinguished from the contribution from the neutral diffuse gas between the large-scale H II regions, i.e. a signature of reionization. I investigated further this impact of clumpy gas around galaxies on the visibility of Lyα emission line from the galaxies. I found that both the galaxy-absorber correlation function and their pairwise velocity field are important to explain the impact on the visibility of Lyα emitting galaxies. Fortunately there is a way to break this degeneracy using surveys of Lyα emitting galaxies. One way is to combine the observations of Lyα luminosity function and Lyα fraction of Lyman break galaxies. As the two statistics of Lyα emitting galaxies are affected differently from small-scale absorbers and large-scale H II regions, the combined analysis informs us about the reionization topology, hence a better constraint on the reionization history.

In Chapter 3, I have studied the role of galaxies and QSOs driving reionization using a large suite of multi-frequency radiative transfer simulations. I have theoretically investigated the properties of the IGM in high-redshift QSO environments where both galaxies and a QSO are thought to exist. Such environments provide a laboratory to study the physical impacts of galaxies and QSOs on the reionization process of the IGM.

I concluded that both the central QSO and surrounding galaxies concertedly control the reionization morphology of hydrogen and helium, and non-linearly impact the thermal structure of the IGM. The morphology, e.g. shape and size, of H II regions is most importantly controlled by the collective contributions from galaxies and the central QSO since the onset of reionization. When a central QSO dominates the total ionizing photons in the environment, the H II region is highly spherical and distinctive from the H II regions created by galaxies. However, since a QSO is likely born in a galaxy overdensity and the cumulative amount of ionizing photons from the galaxies can be comparable to that of the central QSO, the final morphology of the H II region in the QSO environment is a complex product controlled both by galaxies and QSOs. Furthermore, the structures near and beyond the ionization fronts are determined by the spectral shape of the collective radiation field from galaxies and QSOs driving reionization. The X-ray photons from QSOs produce a tail of partially ionized gas ahead of the ionization front. While a UV-obscured QSO can broaden the ionization front, the contribution from other UV sources, either galaxies or unobscured QSO, is sufficient to maintain the sharp ionization front. This implies that there is a physical limitation to extract a detail information about the sources of reionization from the morphology of H II regions, although it helps to narrow down the range of astrophysical parameter space of the sources.

In Chapter 4, I discussed the properties of the radio signals (H I 21cm line and ^{3}He II 3.46cm line) from the intergalactic hydrogen and helium around high redshift QSO environments. The detection of the signals directly probes the physical state of the IGM, and it is a major science driver of ongoing and upcoming radio interferometers such as LOFAR, MWA, PAPER, and SKA. I found that the 21cm signal of H II regions, which appears as coldspots in 21cm tomography, around the galaxy overdensities and QSO environments are likely indistinguishable. It can be nonetheless used to narrow down a parameter space of the source properties responsible for reionization. The direct imaging of H II regions is best utilized by combining infrared facilities such as JWST. The direct search of galaxies and/or QSOs inside the coldspots in a 21cm tomographic image can unambiguously constrain how the sources reionize the surrounding intergalactic gas. The ^{3}He II 3.46cm signal complements the H I 21cm signal. The probe of intergalactic helium allows us to understand the role of QSOs by studying its impact on the helium reionization. The unique helium reionization signature due to QSOs at $z > 6$ could be a smoking gun test about the contribution of QSOs to reionization. The observability of ^{3}He II 3.46cm signal is more complicated than H I 21cm signal because of the subtle physics of ^{3}He II Wouthuysen-Field mechanism involving He II Lyα radiative transfer.

The study of the Epoch of Reionization and Cosmic Dawn is entering the golden era. Upcoming ground-based optical/infrared and radio facilities as well as space-based telescopes provide the observational data to reveal the nature of the high-redshift universe. To fully exploit the data, a suite of cosmological radiative transfer/radiation hydrodynamical simulations plays an important role. A simulation-calibrated measurement using surveys of Lyα emitting galaxies (Chapter 2) exploits the Lyα data to understand the nature of high-redshift star forming galaxies and the intergalactic hydrogen. Based on the detailed study of the role of QSOs during reionization (Chapter 3), one can now be use the muiti-

frequency radiative transfer simulations to understand the QSO absorption spectroscopic data. The method allows us to systematically study both hydrogen and helium reionization on an equal footing. This extends our understanding of the astrophysics of the galaxy-IGM connection from a redshift as high as $z \sim 15$ to $z \sim 2$. The optical/infrared observations and their analysis using simulations/theoretical modelling prepares us to open up the low-frequency radio astronomy, i.e. 21cm cosmology, for more direct study of the EoR and Cosmic Dawn. The deep field observations by the Hubble Space Telescope, both the Ultra Deep Field and the Frontier Fields behind gravitational lenses, offer exquisite data. The additional deep field observations targeting the high-redshift QSOs or gamma-ray bursts will provide unique fields that probe both galaxies and the intergalactic medium at the same time. Because the reionization is a problem of the interplay between galaxies and the intergalactic gas, such new deep fields open up exciting windows for directly studying the heart of the Epoch of Reionization.

Appendix A

Appendix 1

A.1 The mass-weighted neutral fraction in the post-reionized universe

The mass-weighted H I fraction in the post-reionized universe can be estimated from DLA/LLS surveys and Lyα forest observations, which measure the H I column density distribution function. As follows, this quantity can then be converted into the H I fraction embedded as Lyα absorbers, such as DLA, LLS, and diffuse IGM.

The proper number density of H I gas in the universe, $n_{\rm HI}^{prop}(z)$, is expressed as (cf. [295])

$$\begin{aligned} n_{\rm HI}^{prop}(z) &= \int N_{\rm HI} \frac{\partial^2 \mathcal{N}}{\partial N_{\rm HI} \partial z} \left| \frac{dz}{dl_p} \right| dN_{\rm HI}, \\ &= \frac{(1+z)^3 H_0}{c} \int N_{\rm HI} f(N_{\rm HI}, z) dN_{\rm HI}, \end{aligned} \tag{A.1}$$

where l_p is the proper distance, $dl_p/dz = c/H(z)(1+z)$. Therefore, the fraction of neutral hydrogen over the total hydrogen atoms in the entire universe, $\langle f_{\rm HI} \rangle_M$, is given by $\langle f_{\rm HI} \rangle_M = n_{\rm HI}^{prop}(z)/\bar{n}_{\rm H}^{prop}(z)$ [1],

$$\langle f_{\rm HI} \rangle_M = \frac{8\pi G m_H}{3 H_0 c (1-Y) \Omega_b} \int_{N_{\rm HI}^{min}}^{N_{\rm HI}^{max}} N_{\rm HI} f(N_{\rm HI}, z) dN_{\rm HI}, \tag{A.2}$$

where m_H is the mass of a hydrogen atom and $\bar{n}_H^{prop}(z) = \frac{3H_0^2(1-Y)\Omega_b}{8\pi G m_H}(1+z)^3 = 2.057 \times 10^{-7}(1+z)^3 \left(\frac{\Omega_b h^2}{0.023}\right) {\rm cm}^{-3}$ for a helium abundance $Y = 0.25$. The upper and lower limits of the integration specify whether the H I content is embedded in the Lyα forest absorbers ($\log_{10}[N_{\rm HI}/{\rm cm}^{-2}] < 17$), Lyman-limit systems ($17 < \log_{10}[N_{\rm HI}/{\rm cm}^{-2}] < 20.3$), or damped

[1] The fraction of total number of neutral hydrogen, $\mathcal{N}_{\rm H\,I}$, over the total hydrogen atom counts, $\mathcal{N}_{\rm H}$, is given by the mass-weighted neutral fraction $\langle f_{\rm HI} \rangle_M = \mathcal{N}_{\rm H\,I}/\mathcal{N}_{\rm H} = \int x_{\rm HI} n_{\rm H} dV / \int n_{\rm H} dV = \int x_{\rm HI} \rho dV / \int \rho dV$. The volume-weighted and the mass-weighted neutral fraction are identical only for a homogeneous IGM: $\langle f_{\rm HI} \rangle_M = \int x_{\rm HI} \bar{\rho} dV / \int \bar{\rho} dV = \int x_{\rm HI} dV / \int dV = \langle f_{\rm HI} \rangle_V$.

Lyα systems ($20.3 < \log_{10}[N_{\rm HI}/{\rm cm}^{-2}]$). We integrate equation ($A.2$) using the fitting functions to the observed CDDFs, $f(N_{\rm HI}, z)$. We use the CDDF fitting functions from [227] for the Lyα forest absorbers, [347] for the LLS range, and [355] for the DLA range. The observed $f(N_{\rm HI}, z)$ and the various fits are shown in Fig. A.1.

A.2 Effective optical depth of dynamical small-scale absorbers

The opacity from small-scale absorbers is determined by the phase-space distribution function of galaxy-absorber pairs, $f(r_{12}, v_{12}, N_{\rm HI})$, where r_{12} is the comoving separation and v_{12} is the peculiar pairwise radial velocity of pairs.

The line transfer is sensitive to the clustering in total velocity space, $v_c = aHr_{12} + v_{12}$. The probability to find an absorber within v_c and $v_c + dv_c$ and column density $N_{\rm HI}$ and $N_{\rm HI} + dN_{\rm HI}$ is $p(v_c, N_{\rm HI})dv_c dN_{\rm HI}$. Then, the effective optical depth is given by ([331])

$$\tau_{web}^{\rm eff} = \iint p(v_c, N_{\rm HI}) \left[1 - e^{-\tau_{abs}(v_c, N_{\rm HI})}\right] dv_c dN_{\rm HI}. \tag{A.3}$$

$p(v_c, N_{\rm HI})$ is related to the phase-space distribution function of galaxy-absorber pairs through the transformation of variables r_{12}, v_{12} to v_c,

$$\begin{aligned} p(v_c, N_{\rm HI}) &= \iint \delta_D \left[v_c - (aHr_{12} + v_{12})\right] f(r_{12}, v_{12}, N_{\rm HI}) dv_{12} dr_{12} \\ &= \int p_v(v_c - aHr_{12}|r_{12}, N_{\rm HI}) p_r(r_{12}, N_{\rm HI}) dr_{12}, \end{aligned} \tag{A.4}$$

where δ_D is the Dirac delta function. For the second equality, we have used $f(r_{12}, v_{12}, N_{\rm HI}) = p_v(v_{12}|r_{12}, N_{\rm HI}) p_r(r_{12}, N_{\rm HI})$, where $p_v(v_{12}|r_{12}, N_{\rm HI})dv_{12}$ is the conditional probability to find an absorber with peculiar pairwise velocity between v_{12} and $v_{12} + dv_{12}$ at given pair separation r_{12} and column density $N_{\rm HI}$, and $p_r(r_{12}, N_{\rm HI})dr_{12}dN_{\rm HI}$ is the probability to find an absorber in the range r_{12} to $r_{12} + dr_{12}$ and $N_{\rm HI}$ to $N_{\rm HI} + dN_{\rm HI}$. The real-space correlation function $\xi(r_{12}, N_{\rm HI})$ of absorbers around galaxies gives

$$p_r(r_{12}, N_{\rm HI}) = \frac{\partial^2 \mathcal{N}}{\partial N_{\rm HI} \partial z} \left|\frac{dz}{dr}\right| \left[1 + \xi(r_{12}, N_{\rm HI})\right], \tag{A.5}$$

where $|dr/dz| = c/H(z_s)$. Substituting into equation (A.4),

$$p_v(v_c, N_{\rm HI}) = \frac{\partial^2 \mathcal{N}}{\partial N_{\rm HI} \partial z} \left|\frac{dz}{dr}\right| \frac{1}{aH} \left[1 + \xi_v(v_c, N_{\rm HI})\right], \tag{A.6}$$

where we have defined the absorber-galaxy correlation function in velocity space as

$$\begin{aligned} &1 + \xi_v(v_c, N_{\rm HI}) \equiv \\ &aH \int dr_{12} \left[1 + \xi(r_{12}, N_{\rm HI})\right] p_v(v_c - aHr_{12}|r_{12}, N_{\rm HI}). \end{aligned} \tag{A.7}$$

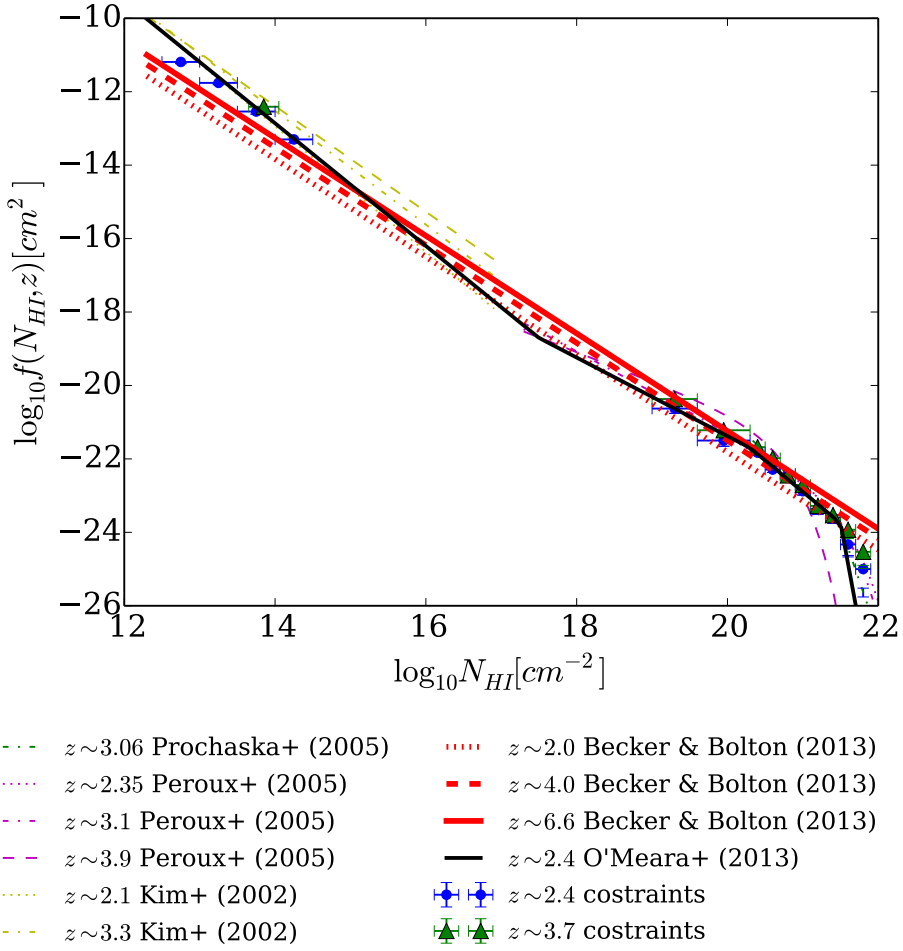

Figure A.1: H I column density distribution function $f(N_{\rm HI}, z)$ at $z \gtrsim 2$. The lines show the fits to the CDDF taken from the literature. The fit by Becker & Bolton (2013) is used to extrapolate to $z > 4$. The points at $z \sim 2.4$ and $z \sim 3.7$ use the compilation of data presented in O'Meara et al. (2013).

Thus, the effective optical depth is

$$\tau_{web}^{\rm eff} = \int dN_{\rm HI} \frac{\partial^2 \mathcal{N}}{\partial N_{\rm HI} \partial z} \left| \frac{dz}{dr} \right| \times \int \frac{dv_c}{aH} (1 + \xi_v(v_c, N_{\rm HI})) \left[1 - e^{-\tau_{abs}(v_c, N_{\rm HI})} \right]. \tag{A.8}$$

All the quantities are evaluated at redshift $z = z_s$. By rearranging we obtain equation (2.14).

In the absence of clustering, $\xi_v = 0$, the effective optical depth (A.8) reduces to the well-known expression for the Poisson-distributed absorbers $\tau_{web}^{\rm eff} = \int dz \int dN_{\rm HI} \left| \frac{dl_p}{dz} \right| \frac{\partial^2 \mathcal{N}}{\partial N_{\rm HI} \partial l_p} (1 - e^{-\tau_{abs}})$ (e.g. [183]).

We show two examples of the velocity-space correlation function ξ_v. For a pure Hubble flow $v_c = aHr_{12}$, $p_v(v_{12}|r_{12}, N_{\rm HI}) = \delta_D(v_{12})$. Thus, $\xi_v(v_c) = \xi\left(r_{12} = \frac{v_c}{aH}\right)$. Furthermore, a Gaussian streaming model is a simple generalization where the conditional pairwise peculiar velocity PDF is modelled as $p_v(v_{12}|r_{12}, N_{\rm HI}) = \frac{1}{\sqrt{2\pi\sigma_{12}^2(r_{12})}} \exp\left[-\frac{(v_{12} - \langle v_{12}(r_{12}) \rangle)^2}{2\sigma_{12}^2(r_{12})} \right]$, where $\langle v_{12}(r_{12}) \rangle$ and $\sigma_{12}(r_{12})$ are the radial pairwise mean peculiar velocity and velocity dispersion, respectively.

A.3 Abundance matching

The abundance matching technique gives a semi-empirical relation between the halo mass and the Lyα luminosity for each f_{duty} as shown in Fig. *A*.2. The red lines are the result of matching the simulated halo mass function at $z = 7$ with the observed $z = 5.7$ Lyα luminosity function ([324]) assuming a duty cycle $f_{duty} = 0.1$ and 1.

Fig. *A*.2 shows that, given a halo mass, a higher duty cycle requires a brighter Lyα luminosity to match the observed $z = 5.7$ Lyα luminosity function, and that a simple functional form, e.g. $L_\alpha \propto M_h,\ M_h^2$, cannot match the semi-empirical relation.

In our model, the intrinsic Lyα luminosity of each galaxy (halo) is assigned according to the $L_\alpha - M_h$ relation with $f_{duty} = 1$ in Fig. *A*.2.

A.4 Lyα RT through the IGM: computing the line-of-sight skewers and optical depth

We compute the Lyα optical depth in the red damping wing as follows. The density, temperature, velocity and local H I fraction fields along skewers originating at the location of halos and parallel to the z-axis are extracted from the hydrodynamical and radiative transfer simulations. To obtain a converged numerical integration of the optical depth, the sampling size of the skewers, δl, must be sufficiently fine. To be on the safe side, the Doppler core of the Voigt line profile should be resolved. In the velocity space this is $\delta v/c = \Delta\nu_D/\nu_\alpha = 4.286 \times 10^{-7} (T/{\rm K})^{1/2}$. Therefore, the velocity space resolution must be

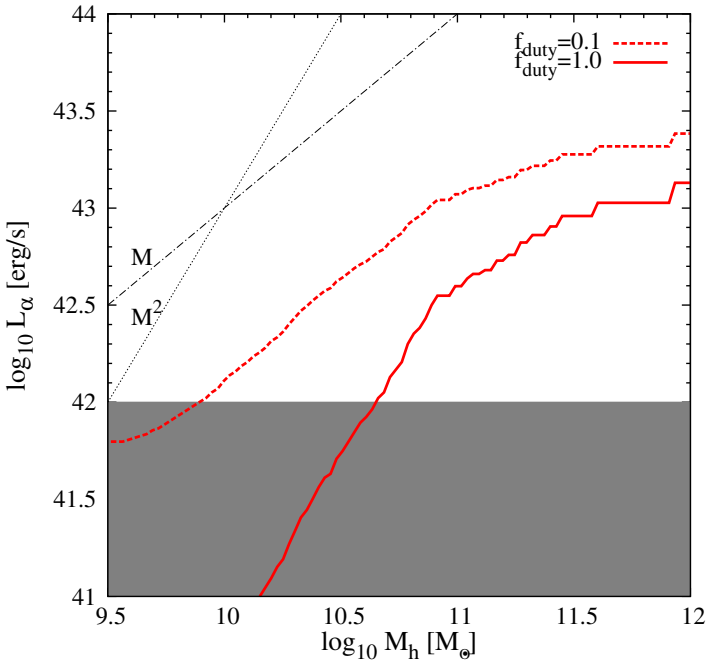

Figure A.2: Semi-empirical relation between the halo mass and the intrinsic Lyα luminosity from the abundance matching technique. The red solid (dashed) line is the result of matching the simulated halo mass function at $z = 7$ with the $z = 5.7$ observed Lyα luminosity function for $f_{duty} = 1.0$ (0.1). The shadowed region indicates the luminosity range below detection limit. The two black dashed lines correspond to $L_\alpha \propto M_h,\ M_h^2$.

$\delta v \approx 0.13(T/1\mathrm{K})^{1/2}$ km/s, which corresponds to a real space resolution of $\delta l = \delta v/H(z_s) \approx 0.17(T/\mathrm{K})^{1/2}$ pkpc at $z_s = 7$ with our cosmological parameters. If this criterion is not met, scattering by Doppler core could be missed. Although the Doppler core scattering is important in low density regions to produce Lyα forest absorption blueward of the rest-frame Lyα line, here we are interested only in the red damping wing and the Lorentz wing scattering. Therefore, a converged evaluation of the optical depth in the red damping wing can still be obtained without strictly meeting this resolution criterion. Nonetheless, the sampling of the line-of-sight skewers must be sufficiently fine, and a sub-sampling within a cell of the cosmological hydrodynamical simulations is required to obtain a convergence in equation (A.9).

To this aim, we have assumed that the density, ionization, temperature and peculiar velocity fields are constant within each cell, while the Hubble flow is allowed to vary. This is required to recover the analytic solution and to obtain a numerically converged optical depth in the limit of homogeneous expanding IGM.

The discretized form of the optical depth is then integrated at each frequency point ν_e using the line-of-sight skewers according to

$$\tau_\alpha(\nu_e) = \sum_{i=1}^{N} \sigma_\alpha n_{\rm HI}(l_i)\varphi_\nu \left[T_i, \nu_e\left(1 - \frac{v_{tot}(l_i)}{c}\right)\right] \delta l. \tag{A.9}$$

The maximum proper length of the line-of-sight skewers influences the far redward optical depth, as a lower length would results in more transmission. We choose the maximum proper length of the skewer to be 12 pMpc. If a skewer exits the simulation box, a random cell in a random face of the box is chosen, and the line-of-sight is followed until the maximum proper length is reached. We have verified that for a homogeneous expanding IGM, the result at $\Delta v \sim 1000$ km/s has a discrepancy of $\sim 8\%$ relative to the analytic solution of the optical depth. Because the IGM will become more ionized as Lyα photons travel through the medium and because we retain the same redshift output to extract the line-of-sight skewers, we choose the maximum length of our skewer samples to be 12 pMpc.

The lower bound of the optical depth integration is chosen to be $300h^{-1}$ckpc. As a reference, the virial radius of a halo with mass M_h is $R_{vir} \approx 78.5(M_h/10^{11}h^{-1}M_\odot)^{1/3}h^{-1}$ckpc, i.e. we exclude from the calculation the gas contained within a halo, as well as all the structures on scales smaller than the Jeans length because they are not well resolved in our simulations.

A.5 Intrinsic Lyα fraction

We write the intrinsic Lyα fraction as $\mathcal{X}^{\rm intr}_{\rm Ly\alpha}(> \mathrm{REW}_{\rm intr}|M_{UV}) = e^{-\mathrm{REW}_{\rm intr}/\mathrm{REW_c}(M_{UV})}$ where $\mathrm{REW_c}(M_{UV})$ is the characteristic REW.

The M_{UV}-dependent model and uncorrelated model differ in their functional form of $\mathrm{REW_c}(M_{UV})$, as the latter assumes a constant $\mathrm{REW_c}(M_{UV}) = 50$ Å, while the former uses the $\mathrm{REW_c}(M_{UV})$ obtained from the best-fit to the Lyα fraction of LBGs observed at $3 < z < 6$ ([425]), i.e. $\mathcal{X}^{\rm intr}_{\rm Ly\alpha}(> \mathrm{REW}|M_{UV}, z = 7) = \mathcal{X}^{3<z<6}_{\rm Ly\alpha}(> \mathrm{REW}|M_{UV})$.

Furthermore, for $P(M_h|M_{UV})$ we assume a one-to-one mapping between UV magnitude and halo mass, i.e. $P(M_h|M_{UV}) = \delta_D(M_h - M_h(M_{UV}))$. The $M_h - M_{UV}$ relation is given by $M_h(M_{UV}) = M_h^* \times 10^{-(M_{UV}-M_{UV}^*)/2.5}$ where $M_h^* = 10^{10}$ $\mathrm{M}_\odot$ and $M_{UV}^* = -19$. We note that this relation tends to assign masses which are typically lower than those derived from observations. For example, $M_h(M_{UV} = -20) = 2.5 \times 10^{10}$ $\mathrm{M}_\odot$, which is much lower than the mass of LBGs hosts inferred from clustering analysis, i.e. $M_h \sim 3 \times 10^{11} - 10^{12}$ $\mathrm{M}_\odot$ (e.g. [219]). Since we expect the dependence of $\mathcal{T}_{IGM}$ on halo mass to extends in the range $11 < \log_{10} M_h/\mathrm{M}_\odot < 12$, we assume the sampling of the $\mathcal{T}_{IGM} - M_h$ relation at low mass haloes to mimic the realistic host halo mass of observed LBGs.

Bibliography

[1] T. Abel und M.G. Haehnelt, *ApJ* **520** (1999), L13.

[2] T. Abel, M.L. Norman und P. Madau, *ApJ* **523** (1999), 66.

[3] K.L. Adelberger, C.C. Steidel, M. Pettini, A.E. Shapley, N.A. Reddy und D.K. Erb, *ApJ* **619** (2005), 697.

[4] K.L. Adelberger, C.C. Steidel, A.E. Shapley und M. Pettini, *ApJ* **584** (2003), 45.

[5] K. Ahn, H. Xu, M.L. Norman, M.A. Alvarez und J.H. Wise, *ApJ* **802** (2015), 8.

[6] Z.S. Ali, A.R. Parsons, H. Zheng und et al., *ApJ* **809** (2015), 61.

[7] M.A. Alvarez und T. Abel, *MNRAS* **380** (2007), L30.

[8] L. Anderson, É. Aubourg, S. Bailey und et al., *MNRAS* **441** (2014), 24.

[9] M. Ando, K. Ohta, I. Iwata, M. Akiyama, K. Aoki und N. Tamura, *ApJ* **645** (2006), L9.

[10] S. Baek, B. Semelin, P. Di Matteo, Y. Revaz und F. Combes, *A&Ap* **523** (2010), A4.

[11] J.S. Bagla und A. Loeb, *ArXiv e-prints* (2009).

[12] R. Barkana und A. Loeb, *ApJ* **601** (2004), 64.

[13] R. Barkana und A. Loeb, *ApJ* **609** (2004), 474.

[14] R. Barkana und A. Loeb, *ApJ* **626** (2005), 1.

[15] L.A. Barnes, T. Garel und G.G. Kacprzak, *PASP* **126** (2014), 969.

[16] R.L. Barone-Nugent, M. Trenti, J.S.B. Wyithe, R.J. Bouwens, P.A. Oesch, G.D. Illingworth, C.M. Carollo, J. Su, M. Stiavelli, I. Labbe und P.G. van Dokkum, *ApJ* **793** (2014), 17.

[17] A. Bauer, V. Springel, M. Vogelsberger, S. Genel, P. Torrey, D. Sijacki, D. Nelson und L. Hernquist, *MNRAS* **453** (2015), 3593.

[18] G.D. Becker und J.S. Bolton, *MNRAS* **436** (2013), 1023.

[19] G.D. Becker, J.S. Bolton, M.G. Haehnelt und W.L.W. Sargent, *MNRAS* **410** (2011), 1096.

[20] G.D. Becker, J.S. Bolton und A. Lidz, *PASA* **32** (2015), e045.

[21] G.D. Becker, J.S. Bolton und A. Lidz, *ArXiv e-prints* (2015).

[22] G.D. Becker, J.S. Bolton, P. Madau, M. Pettini, E.V. Ryan-Weber und B.P. Venemans, *MNRAS* **447** (2015), 3402.

[23] G.D. Becker, M. Rauch und W.L.W. Sargent, *ApJ* **662** (2007), 72.

[24] G.F. Benedict, B.E. McArthur, M.W. Feast, T.G. Barnes, T.E. Harrison, R.J. Patterson, J.W. Menzies, J.L. Bean und W.L. Freedman, *AJ* **133** (2007), 1810.

[25] A.J. Benson, *Phys. Rep.* **495** (2010), 33.

[26] E. Bertschinger, *ArXiv Astrophysics e-prints* (1995).

[27] E. Bertschinger: *Cosmological Dynamics*. *Cosmological Dynamics*, In *Cosmology and Large Scale Structure*, herausgegeben von R. Schaeffer, J. Silk, M. Spiro und J. Zinn-Justin. (Januar 1996) Seite 273.

[28] E. Bertschinger, *ARA&A* **36** (1998), 599.

[29] P.N. Best und t. LOFAR-UK Consortium, *ArXiv e-prints* (2008).

[30] F. Beutler, C. Blake, M. Colless, D.H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders und F. Watson, *MNRAS* **416** (2011), 3017.

[31] R. Bielby, M.D. Hill, T. Shanks und et al., *MNRAS* **430** (2013), 425.

[32] R.M. Bielby, P. Tummuangpak, T. Shanks und et al., *MNRAS* **456** (2016), 4061.

[33] Y. Birnboim und A. Dekel, *MNRAS* **345** (2003), 349.

[34] C. Blake, E.A. Kazin, F. Beutler und et al., *MNRAS* **418** (2011), 1707.

[35] G.A. Blanc, J.J. Adams, K. Gebhardt und et al., *ApJ* **736** (2011), 31.

[36] M.R. Blanton und J. Moustakas, *ARA&A* **47** (2009), 159.

[37] E. Boera, M.T. Murphy, G.D. Becker und J.S. Bolton, *MNRAS* **441** (2014), 1916.

[38] E. Boera, M.T. Murphy, G.D. Becker und J.S. Bolton, *MNRAS* **456** (2016), L79.

[39] J.S. Bolton, G.D. Becker, M.G. Haehnelt und M. Viel, *MNRAS* **438** (2014), 2499.

[40] J.S. Bolton, G.D. Becker, S. Raskutti, J.S.B. Wyithe, M.G. Haehnelt und W.L.W. Sargent, *MNRAS* **419** (2012), 2880.

[41] J.S. Bolton, G.D. Becker, J.S.B. Wyithe, M.G. Haehnelt und W.L.W. Sargent, *MNRAS* **406** (2010), 612.

[42] J.S. Bolton und M.G. Haehnelt, *MNRAS* **374** (2007), 493.

[43] J.S. Bolton und M.G. Haehnelt, *MNRAS* **382** (2007), 325.

[44] J.S. Bolton und M.G. Haehnelt, *MNRAS* **429** (2013), 1695.

[45] J.S. Bolton, M.G. Haehnelt, M. Viel und R.F. Carswell, *MNRAS* **366** (2006), 1378.

[46] J.S. Bolton, M.G. Haehnelt, M. Viel und V. Springel, *MNRAS* **357** (2005), 1178.

[47] J.S. Bolton, M.G. Haehnelt, S.J. Warren, P.C. Hewett, D.J. Mortlock, B.P. Venemans, R.G. McMahon und C. Simpson, *MNRAS* **416** (2011), L70.

[48] J.S. Bolton, M. Viel, T.S. Kim, M.G. Haehnelt und R.F. Carswell, *MNRAS* **386** (2008), 1131.

[49] J.R. Bond, G. Efstathiou und M. Tegmark, *MNRAS* **291** (1997), L33.

[50] S.E.I. Bosman und G.D. Becker, *MNRAS* **452** (2015), 1105.

[51] F. Bournaud, L.C. Powell, D. Chapon und R. Teyssier: *Star formation in galaxy mergers: ISM turbulence, dense gas excess, and scaling relations for disks and starbusts. Star formation in galaxy mergers: ISM turbulence, dense gas excess, and scaling relations for disks and starbusts*, In *Astrophysical Dynamics: From Stars to Galaxies*, herausgegeben von N. H. Brummell, A. S. Brun, M. S. Miesch und Y. Ponty, Band 271 von *IAU Symposium*. (August 2011) Seiten 160–169.

[52] R. Bouwens: *High-Redshift Galaxy Surveys and the Reionization of the Universe. High-Redshift Galaxy Surveys and the Reionization of the Universe*, In *Astrophysics and Space Science Library*, herausgegeben von A. Mesinger, Band 423 von *Astrophysics and Space Science Library*. (2016) Seite 111.

[53] R.J. Bouwens, G.D. Illingworth, P.A. Oesch, M. Trenti, I. Labbé, L. Bradley, M. Carollo, P.G. van Dokkum, V. Gonzalez, B. Holwerda, M. Franx, L. Spitler, R. Smit und D. Magee, *ApJ* **803** (2015), 34.

[54] R.A.A. Bowler, J.S. Dunlop, R.J. McLure, H.J. McCracken, B. Milvang-Jensen, H. Furusawa, Y. Taniguchi, O. Le Fèvre, J.P.U. Fynbo, M.J. Jarvis und B. Häußler, *MNRAS* **452** (2015), 1817.

[55] R.A.A. Bowler, J.S. Dunlop, R.J. McLure, A.B. Rogers, H.J. McCracken, B. Milvang-Jensen, H. Furusawa, J.P.U. Fynbo, Y. Taniguchi, J. Afonso, M.N. Bremer und O. Le Fèvre, *MNRAS* **440** (2014), 2810.

[56] J.D. Bowman und A.E.E. Rogers, *Nature* **468** (2010), 796.

[57] J.D. Bowman, A.E.E. Rogers und J.N. Hewitt, *ApJ* **676** (2008), 1.

[58] V. Bromm und N. Yoshida, *ARA&A* **49** (2011), 373.

[59] G.L. Bryan und M.E. Machacek, *ApJ* **534** (2000), 57.

[60] G.L. Bryan, M.L. Norman, B.W. O'Shea und et al., *ApJS* **211** (2014), 19.

[61] N.G. Busca, T. Delubac, J. Rich und et al., *A&Ap* **552** (2013), A96.

[62] F. Calura, E. Tescari, V. D'Odorico, M. Viel, S. Cristiani, T.S. Kim und J.S. Bolton, *MNRAS* **422** (2012), 3019.

[63] A.P. Calverley, G.D. Becker, M.G. Haehnelt und J.S. Bolton, *MNRAS* **412** (2011), 2543.

[64] J. Caruana, A.J. Bunker, S.M. Wilkins, E.R. Stanway, S. Lorenzoni, M.J. Jarvis und H. Ebert, *MNRAS* **443** (2014), 2831.

[65] P. Cassata, O. Le Fèvre, B. Garilli und et al., *A&Ap* **525** (2011), A143.

[66] P. Cassata, L.A.M. Tasca, O. Le Fèvre und et al., *A&Ap* **573** (2015), A24.

[67] M. Castellano, P. Dayal, L. Pentericci und e al., *ApJ* **818** (2016), L3.

[68] J.I. Castor: *Radiation Hydrodynamics*, November 2004.

[69] R. Cen, *ApJS* **78** (1992), 341.

[70] R. Cen und Z. Haiman, *ApJ* **542** (2000), L75.

[71] R. Cen, J. Miralda-Escudé, J.P. Ostriker und M. Rauch, *ApJ* **437** (1994), L9.

[72] E. Chapman, F.B. Abdalla, J. Bobin und J.L.e.a. Starck, *MNRAS* **429** (2013), 165.

[73] J. Chardin, M.G. Haehnelt, D. Aubert und E. Puchwein, *ArXiv e-prints* (2015).

[74] X. Chen und J. Miralda-Escudé, *ApJ* **684** (2008), 18.

[75] R. Chornock, E. Berger, D.B. Fox, W. Fong, T. Laskar und K.C. Roth, *ArXiv e-prints* (2014).

[76] R. Chornock, E. Berger, D.B. Fox, R. Lunnan, M.R. Drout, W.f. Fong, T. Laskar und K.C. Roth, *ApJ* **774** (2013), 26.

[77] T.R. Choudhury, E. Puchwein, M.G. Haehnelt und J.S. Bolton, *MNRAS* **452** (2015), 261.

[78] L. Chuzhoy, M.A. Alvarez und P.R. Shapiro, *ApJ* **648** (2006), L1.

[79] L. Chuzhoy und P.R. Shapiro, *ApJ* **651** (2006), 1.

[80] B. Ciardi, J.S. Bolton, A. Maselli und L. Graziani, *MNRAS* **423** (2012), 558.

[81] B. Ciardi, A. Ferrara, S. Marri und G. Raimondo, *MNRAS* **324** (2001), 381.

[82] B. Ciardi, A. Ferrara und S.D.M. White, *MNRAS* **344** (2003), L7.

[83] L. Ciotti und J.P. Ostriker, *ApJ* **551** (2001), 131.

[84] S. Cole, W.J. Percival, J.A. Peacock und et al., *MNRAS* **362** (2005), 505.

[85] M. Colless, G. Dalton, S. Maddox und et al., *MNRAS* **328** (2001), 1039.

[86] M. Compostella, S. Cantalupo und C. Porciani, *MNRAS* **435** (2013), 3169.

[87] M. Compostella, S. Cantalupo und C. Porciani, *MNRAS* **445** (2014), 4186.

[88] J. Cooke, *ApJ* **704** (2009), L62.

[89] J. Cooke, Y. Omori und E.V. Ryan-Weber, *MNRAS* **433** (2013), 2122.

[90] J. Cooke, A.M. Wolfe, E. Gawiser und J.X. Prochaska, *ApJ* **636** (2006), L9.

[91] J. Cooke, A.M. Wolfe, E. Gawiser und J.X. Prochaska, *ApJ* **652** (2006), 994.

[92] N.H.M. Crighton, R. Bielby, T. Shanks und et al., *MNRAS* **414** (2011), 28.

[93] N.H.M. Crighton, M.T. Murphy, J.X. Prochaska und et al., *MNRAS* **452** (2015), 217.

[94] R.A.C. Croft, *ApJ* **610** (2004), 642.

[95] R.A.C. Croft, L. Hernquist, V. Springel, M. Westover und M. White, *ApJ* **580** (2002), 634.

[96] R.A.C. Croft, D.H. Weinberg, M. Bolte, S. Burles, L. Hernquist, N. Katz, D. Kirkman und D. Tytler, *ApJ* **581** (2002), 20.

[97] R.A.C. Croft, D.H. Weinberg, M. Pettini, L. Hernquist und N. Katz, *ApJ* **520** (1999), 1.

[98] A. Cucchiara, A.J. Levan, D.B. Fox und et al., *ApJ* **736** (2011), 7.

[99] E. Curtis-Lake, R.J. McLure, H.J. Pearce, J.S. Dunlop, M. Cirasuolo, D.P. Stark, O. Almaini, E.J. Bradshaw, R. Chuter, S. Foucaud und W.G. Hartley, *MNRAS* **422** (2012), 1425.

[100] A. Dalgarno, M. Yan und W. Liu, *ApJS* **125** (1999), 237.

[101] A. Dall'Aglio, L. Wisotzki und G. Worseck, *A&Ap* **491** (2008), 465.

[102] A. D'Aloisio, M. McQuinn und H. Trac, *ApJ* **813** (2015), L38.

[103] K.K. Datta, M.M. Friedrich und G.e.a. Mellema, *MNRAS* **424** (2012), 762.

[104] M. Davis und P.J.E. Peebles, *ApJ* **267** (1983), 465.

[105] K.S. Dawson, D.J. Schlegel, C.P. Ahn und et al., *AJ* **145** (2013), 10.

[106] P. Dayal und A. Ferrara, *MNRAS* **421** (2012), 2568.

[107] P. Dayal, A. Maselli und A. Ferrara, *MNRAS* **410** (2011), 830.

[108] S. de Barros, D. Schaerer und D.P. Stark, *A&Ap* **563** (2014), A81.

[109] S. Deguchi und W.D. Watson, *ApJ* **290** (1985), 578.

[110] A. Dekel, Y. Birnboim, G. Engel, J. Freundlich, T. Goerdt, M. Mumcuoglu, E. Neistein, C. Pichon, R. Teyssier und E. Zinger, *Nature* **457** (2009), 451.

[111] T. Delubac, J.E. Bautista, N.G. Busca und et al., *A&Ap* **574** (2015), A59.

[112] V. Desjacques, A. Nusser, M.G. Haehnelt und F. Stoehr, *MNRAS* **350** (2004), 879.

[113] T. Di Matteo, N. Khandai, C. DeGraf, Y. Feng, R.A.C. Croft, J. Lopez und V. Springel, *ApJ* **745** (2012), L29.

[114] C.G. Díaz, Y. Koyama, E.V. Ryan-Weber, J. Cooke, M. Ouchi, K. Shimasaku und F. Nakata, *MNRAS* **442** (2014), 946.

[115] C.G. Díaz, E.V. Ryan-Weber, J. Cooke, Y. Koyama und M. Ouchi, *MNRAS* **448** (2015), 1240.

[116] M. Dijkstra, *PASA* **31** (2014), e040.

[117] M. Dijkstra: *Constraining Reionization with Lyα Emitting Galaxies. Constraining Reionization with Lyα Emitting Galaxies*, In *Astrophysics and Space Science Library*, herausgegeben von A. Mesinger, Band 423 von *Astrophysics and Space Science Library*. (2016) Seite 145.

[118] M. Dijkstra, Z. Haiman und A. Loeb, *ApJ* **613** (2004), 646.

[119] M. Dijkstra, Z. Haiman und M. Spaans, *ApJ* **649** (2006), 14.

[120] M. Dijkstra, A. Lidz und J.S.B. Wyithe, *MNRAS* **377** (2007), 1175.

[121] M. Dijkstra, A. Mesinger und J.S.B. Wyithe, *MNRAS* **414** (2011), 2139.

[122] M. Dijkstra und J.S.B. Wyithe, *MNRAS* **408** (2010), 352.

[123] M. Dijkstra und J.S.B. Wyithe, *MNRAS* **419** (2012), 3181.

[124] M. Dijkstra, S. Wyithe, Z. Haiman, A. Mesinger und L. Pentericci, *MNRAS* **440** (2014), 3309.

[125] J.S. Dillon, A. Liu, C.L. Williams und et al., *Phys. Rev. D* **89** (2014), 023002.

[126] J.S. Dillon, A.R. Neben, J.N. Hewitt und et al., *Phys. Rev. D* **91** (2015), 123011.

[127] S. Dodelson: *Modern cosmology*, 2003.

[128] B.T. Draine: *Physics of the Interstellar and Intergalactic Medium*, 2011.

[129] A. Dressler, *ApJ* **236** (1980), 351.

[130] A. Dressler, A. Henry, C.L. Martin, M. Sawicki, P. McCarthy und E. Villaneuva, *ApJ* **806** (2015), 19.

[131] J.S. Dunlop: *Observing the First Galaxies. Observing the First Galaxies*, In *The First Galaxies*, herausgegeben von T. Wiklind, B. Mobasher und V. Bromm, Band 396 von *Astrophysics and Space Science Library*. (2013) Seite 223.

[132] G. Efstathiou: *Observations of Large-Scale Structure in the Universe. Observations of Large-Scale Structure in the Universe*, In *Cosmology and Large Scale Structure*, herausgegeben von R. Schaeffer, J. Silk, M. Spiro und J. Zinn-Justin. (Januar 1996) Seite 133.

[133] G. Efstathiou, *MNRAS* **440** (2014), 1138.

[134] D.J. Eisenstein und P. Hut, *ApJ* **498** (1998), 137.

[135] D.J. Eisenstein, D.H. Weinberg, E. Agol, H. Aihara, C. Allende Prieto, S.F. Anderson, J.A. Arns, É. Aubourg, S. Bailey, E. Balbinot und et al., *AJ* **142** (2011), 72.

[136] D.J. Eisenstein, I. Zehavi, D.W. Hogg und et al., *ApJ* **633** (2005), 560.

[137] R.S. Ellis. *Observations of the High Redshift Universe*, 2008. Seiten 259–364.

[138] R.S. Ellis, R.J. McLure, J.S. Dunlop, B.E. Robertson, Y. Ono, M.A. Schenker, A. Koekemoer, R.A.A. Bowler, M. Ouchi, A.B. Rogers, E. Curtis-Lake, E. Schneider, S. Charlot, D.P. Stark, S.R. Furlanetto und M. Cirasuolo, *ApJ* **763** (2013), L7.

[139] X. Fan, C.L. Carilli und B. Keating, *ARA&A* **44** (2006), 415.

[140] X. Fan, M.A. Strauss, R.H. Becker, R.L. White, J.E. Gunn, G.R. Knapp, G.T. Richards, D.P. Schneider, J. Brinkmann und M. Fukugita, *AJ* **132** (2006), 117.

[141] C.A. Faucher-Giguère, A. Lidz, L. Hernquist und M. Zaldarriaga, *ApJ* **682** (2008), L9.

[142] C.A. Faucher-Giguère, A. Lidz, M. Zaldarriaga und L. Hernquist, *ApJ* **703** (2009), 1416.

[143] Y. Feng, R.A.C. Croft, T. Di Matteo und N. Khandai, *MNRAS* **429** (2013), 1554.

[144] Y. Feng, T. Di-Matteo, R.A. Croft, S. Bird, N. Battaglia und S. Wilkins, *ArXiv e-prints* (2015).

[145] A. Fialkov und R. Barkana, *MNRAS* **445** (2014), 213.

[146] G.B. Field, *Proceedings of the IRE* **46** (1958), 240.

[147] G.B. Field, *ApJ* **129** (1959), 536.

[148] S.L. Finkelstein, *ArXiv e-prints* (2015).

[149] K. Finlator, F. Özel und R. Davé, *MNRAS* **393** (2009), 1090.

[150] K. Finlator, R. Thompson, S. Huang, R. Davé, E. Zackrisson und B.D. Oppenheimer, *MNRAS* **447** (2015), 2526.

[151] A. Fontana, E. Vanzella, L. Pentericci, M. Castellano, M. Giavalisco, A. Grazian, K. Boutsia, S. Cristiani, M. Dickinson, E. Giallongo, R. Maiolino, A. Moorwood und P. Santini, *ApJ* **725** (2010), L205.

[152] W.L. Freedman und B.F. Madore, *ARA&A* **48** (2010), 673.

[153] W.L. Freedman, B.F. Madore, B.K. Gibson, L. Ferrarese, D.D. Kelson, S. Sakai, J.R. Mould, R.C. Kennicutt, Jr., H.C. Ford, J.A. Graham, J.P. Huchra, S.M.G. Hughes, G.D. Illingworth, L.M. Macri und P.B. Stetson, *ApJ* **553** (2001), 47.

[154] B. Fryxell, K. Olson, P. Ricker, F.X. Timmes, M. Zingale, D.Q. Lamb, P. MacNeice, R. Rosner, J.W. Truran und H. Tufo, *ApJS* **131** (2000), 273.

[155] M. Fukugita und P.J.E. Peebles, *ApJ* **616** (2004), 643.

[156] S.R. Furlanetto, Z. Haiman und S.P. Oh, *ApJ* **686** (2008), 25.

[157] S.R. Furlanetto und A. Mesinger, *MNRAS* **394** (2009), 1667.

[158] S.R. Furlanetto und S.P. Oh, *MNRAS* **363** (2005), 1031.

[159] S.R. Furlanetto, S.P. Oh und F.H. Briggs, *Phys. Rep.* **433** (2006), 181.

[160] S.R. Furlanetto und S.J. Stoever, *MNRAS* **404** (2010), 1869.

[161] S.R. Furlanetto, M. Zaldarriaga und L. Hernquist, *ApJ* **613** (2004), 1.

[162] G. Galilei: *Dialogue Concerning the Two Chief World Systems*. Princeton University Press, Princeton, 1632.

[163] S. Gallerani, T.R. Choudhury und A. Ferrara, *MNRAS* **370** (2006), 1401.

[164] S. Gallerani, A. Ferrara, X. Fan und T.R. Choudhury, *MNRAS* **386** (2008), 359.

[165] S. Gallerani, R. Salvaterra, A. Ferrara und T.R. Choudhury, *MNRAS* **388** (2008), L84.

[166] L. Gao, S.D.M. White, A. Jenkins, C.S. Frenk und V. Springel, *MNRAS* **363** (2005), 379.

[167] A. Garzilli, J.S. Bolton, T.S. Kim, S. Leach und M. Viel, *MNRAS* **424** (2012), 1723.

[168] E. Gawiser, H. Francke, K. Lai, K. Schawinski, C. Gronwall, R. Ciardullo, R. Quadri, A. Orsi, L.F. Barrientos, G.A. Blanc, G. Fazio, J.J. Feldmeier, J.s. Huang, L. Infante, P. Lira, N. Padilla, E.N. Taylor, E. Treister, C.M. Urry, P.G. van Dokkum und S.N. Virani, *ApJ* **671** (2007), 278.

[169] P.M. Geil und J.S.B. Wyithe, *MNRAS* **386** (2008), 1683.

[170] E. Giallongo, A. Grazian, F. Fiore, A. Fontana und et. al., *A&Ap* **578** (2015), A83.

[171] E. Glikman, S.G. Djorgovski, D. Stern, A. Dey, B.T. Jannuzi und K.S. Lee, *ApJ* **728** (2011), L26.

[172] N.Y. Gnedin, *ApJ* **535** (2000), 530.

[173] N.Y. Gnedin, *ApJ* **793** (2014), 29.

[174] N.Y. Gnedin und T. Abel, *New Astro.* **6** (2001), 437.

[175] N.Y. Gnedin, A.V. Kravtsov und H.W. Chen, *ApJ* **672** (2008), 765.

[176] A. Goobar und B. Leibundgut, *Annual Review of Nuclear and Particle Science* **61** (2011), 251.

[177] L. Graziani, A. Maselli und B. Ciardi, *MNRAS* **431** (2013), 722.

[178] B. Greig, J.S. Bolton und J.S.B. Wyithe, *MNRAS* **447** (2015), 2503.

[179] J. Greiner, T. Krühler, J.P.U. Fynbo und et al., *ApJ* **693** (2009), 1610.

[180] M. Gronke, P. Bull und M. Dijkstra, *ArXiv e-prints* (2015).

[181] M. Gronke, M. Dijkstra, M. Trenti und S. Wyithe, *MNRAS* **449** (2015), 1284.

[182] J.E. Gunn und B.A. Peterson, *ApJ* **142** (1965), 1633.

[183] F. Haardt und P. Madau, *ApJ* **461** (1996), 20.

[184] F. Haardt und P. Madau, *ApJ* **746** (2012), 125.

[185] F. Haardt und R. Salvaterra, *A&Ap* **575** (2015), L16.

[186] Y. Harikane, M. Ouchi, Y. Ono und et al., *ArXiv e-prints* (2015).

[187] O.E. Hartoog, D. Malesani, J.P.U. Fynbo und et al., *A&Ap* **580** (2015), A139.

[188] E. Hawkins, S. Maddox, S. Cole und et al., *MNRAS* **346** (2003), 78.

[189] M. Hayes, D. Schaerer, G. Östlin, J.M. Mas-Hesse, H. Atek und D. Kunth, *ApJ* **730** (2011), 8.

[190] L. Hernquist, N. Katz, D.H. Weinberg und J. Miralda-Escudé, *ApJ* **457** (1996), L51.

[191] H. Hildebrandt, J. Pielorz, T. Erben, L. van Waerbeke, P. Simon und P. Capak, *A&Ap* **498** (2009), 725.

[192] W. Hillebrandt und J.C. Niemeyer, *ARA&A* **38** (2000), 191.

[193] G. Hinshaw, D. Larson, E. Komatsu und et al., *ApJS* **208** (2013), 19.

[194] R.W. Hockney und J.W. Eastwood: *Computer Simulation Using Particles*, 1981.

[195] P.F. Hopkins, L. Hernquist, T.J. Cox, T. Di Matteo, B. Robertson und V. Springel, *ApJS* **163** (2006), 1.

[196] P.F. Hopkins, D. Kereš, J. Oñorbe, C.A. Faucher-Giguère, E. Quataert, N. Murray und J.S. Bullock, *MNRAS* **445** (2014), 581.

[197] P.F. Hopkins, E. Quataert und N. Murray, *MNRAS* **417** (2011), 950.

[198] P.F. Hopkins, E. Quataert und N. Murray, *MNRAS* **421** (2012), 3522.

[199] E.M. Hu, L.L. Cowie, A.J. Barger, P. Capak, Y. Kakazu und L. Trouille, *ApJ* **725** (2010), 394.

[200] W. Hu und S. Dodelson, *ARA&A* **40** (2002), 171.

[201] E. Hubble, *Proceedings of the National Academy of Science* **15** (1929), 168.

[202] E. Hubble und M.L. Humason, *ApJ* **74** (1931), 43.

[203] L. Hui und N.Y. Gnedin, *MNRAS* **292** (1997), 27.

[204] K. Husband, M.N. Bremer, E.R. Stanway, L.J.M. Davies, M.D. Lehnert und L.S. Douglas, *MNRAS* **432** (2013), 2869.

[205] A. Hutter, P. Dayal und V. Müller, *MNRAS* **450** (2015), 4025.

[206] A. Hutter, P. Dayal, A.M. Partl und V. Müller, *MNRAS* **441** (2014), 2861.

[207] I.T. Iliev, B. Ciardi, M.A. Alvarez, A. Maselli, A. Ferrara, N.Y. Gnedin, G. Mellema, T. Nakamoto, M.L. Norman, A.O. Razoumov, E.J. Rijkhorst, J. Ritzerveld, P.R. Shapiro, H. Susa, M. Umemura und D.J. Whalen, *MNRAS* **371** (2006), 1057.

[208] I.T. Iliev, G. Mellema, U.L. Pen, H. Merz, P.R. Shapiro und M.A. Alvarez, *MNRAS* **369** (2006), 1625.

[209] I.T. Iliev, G. Mellema, P.R. Shapiro und U.L. Pen, *MNRAS* **376** (2007), 534.

[210] A. Jeeson-Daniel, B. Ciardi, U. Maio, M. Pierleoni, M. Dijkstra und A. Maselli, *MNRAS* **424** (2012), 2193.

[211] V. Jelić, A.G. de Bruyn, M. Mevius, F.B. Abdalla und et al., *A&Ap* **568** (2014), A101.

[212] V. Jelić, S. Zaroubi, P. Labropoulos, G. Bernardi, A.G. de Bruyn und L.V.E. Koopmans, *MNRAS* **409** (2010), 1647.

[213] H. Jensen, M. Hayes, I.T. Iliev, P. Laursen, G. Mellema und E. Zackrisson, *MNRAS* **444** (2014), 2114.

[214] H. Jensen, P. Laursen, G. Mellema, I.T. Iliev, J. Sommer-Larsen und P.R. Shapiro, *MNRAS* **428** (2013), 1366.

[215] L. Jiang, E. Egami, M. Mechtley und et al., *ApJ* **772** (2013), 99.

[216] T.A. Jones, R.S. Ellis, M.A. Schenker und D.P. Stark, *ApJ* **779** (2013), 52.

[217] K. Kakiichi, M. Dijkstra, B. Ciardi und L. Graziani, *ArXiv e-prints* (2015).

[218] N. Kashikawa, K. Shimasaku, M.A. Malkan und et al., *ApJ* **648** (2006), 7.

[219] N. Kashikawa, M. Yoshida, K. Shimasaku und et al., *ApJ* **637** (2006), 631.

[220] N. Kawai, G. Kosugi, K. Aoki und et al., *Nature* **440** (2006), 184.

[221] L.C. Keating, M.G. Haehnelt, S. Cantalupo und E. Puchwein, *MNRAS* **454** (2015), 681.

[222] R.C. Kennicutt, Jr., *ApJ* **498** (1998), 541.

[223] R.C. Kennicutt, Jr., W.L. Freedman und J.R. Mould, *AJ* **110** (1995), 1476.

[224] D. Kereš, N. Katz, D.H. Weinberg und R. Davé, *MNRAS* **363** (2005), 2.

[225] S. Kim, M. Stiavelli, M. Trenti, C.M. Pavlovsky, S.G. Djorgovski, C. Scarlata, D. Stern, A. Mahabal, D. Thompson, M. Dickinson, N. Panagia und G. Meylan, *ApJ* **695** (2009), 809.

[226] T.S. Kim, J.S. Bolton, M. Viel, M.G. Haehnelt und R.F. Carswell, *MNRAS* **382** (2007), 1657.

[227] T.S. Kim, R.F. Carswell, S. Cristiani, S. D'Odorico und E. Giallongo, *MNRAS* **335** (2002), 555.

[228] T.S. Kim, A.M. Partl, R.F. Carswell und V. Müller, *A&Ap* **552** (2013), A77.

[229] T.S. Kim, M. Viel, M.G. Haehnelt, R.F. Carswell und S. Cristiani, *MNRAS* **347** (2004), 355.

[230] T. Kimm und R. Cen, *ApJ* **788** (2014), 121.

[231] R. Kippenhahn, A. Weigert und A. Weiss: *Stellar Structure and Evolution*, 2012.

[232] E. Komatsu, C.L. Bennett, C. Barnes und et al., *Progress of Theoretical and Experimental Physics* **2014** (2014), 06B102.

[233] E. Komatsu, K.M. Smith, J. Dunkley und et al., *ApJS* **192** (2011), 18.

[234] A. Konno, M. Ouchi, Y. Ono, K. Shimasaku, T. Shibuya, H. Furusawa, K. Nakajima, Y. Naito, R. Momose, S. Yuma und M. Iye, *ApJ* **797** (2014), 16.

[235] K. Kovač, R.S. Somerville, J.E. Rhoads, S. Malhotra und J. Wang, *ApJ* **668** (2007), 15.

[236] R.H. Kramer und Z. Haiman, *MNRAS* **385** (2008), 1561.

[237] A.V. Kravtsov: *High-resolution simulations of structure formation in the universe*. NEW MEXICO STATE UNIVERSITY, Dissertation, 1999.

[238] M.R. Krumholz: *Star Formation in Molecular Clouds. Star Formation in Molecular Clouds*, In *American Institute of Physics Conference Series*, herausgegeben von E. Telles, R. Dupke und D. Lazzaro, Band 1386 von *American Institute of Physics Conference Series*. (September 2011) Seiten 9–57.

[239] M.R. Krumholz und J.C. Tan, *ApJ* **654** (2007), 304.

[240] M. Kuhlen und C.A. Faucher-Giguère, *MNRAS* **423** (2012), 862.

[241] F. La Franca, F. Fiore, A. Comastri und et al., *ApJ* **635** (2005), 864.

[242] C. Lacey und S. Cole, *MNRAS* **262** (1993), 627.

[243] O. Lahav und Y. Suto, *Living Reviews in Relativity* **7** (2004).

[244] K. Lai, A. Lidz, L. Hernquist und M. Zaldarriaga, *ApJ* **644** (2006), 61.

[245] P. Laursen, J. Sommer-Larsen und A.O. Razoumov, *ApJ* **728** (2011), 52.

[246] O. Le Fèvre, L.A.M. Tasca, P. Cassata und et al., *A&Ap* **576** (2015), A79.

[247] K.G. Lee, J.F. Hennawi, D.N. Spergel, D.H. Weinberg, D.W. Hogg, M. Viel, J.S. Bolton, S. Bailey, M.M. Pieri, W. Carithers, D.J. Schlegel, B. Lundgren, N. Palanque-Delabrouille, N. Suzuki, D.P. Schneider und C. Yèche, *ApJ* **799** (2015), 196.

[248] K.S. Lee, M. Giavalisco, O.Y. Gnedin, R.S. Somerville, H.C. Ferguson, M. Dickinson und M. Ouchi, *ApJ* **642** (2006), 63.

[249] T.S. Lee, K. Nagamine, L. Hernquist und V. Springel, *MNRAS* **411** (2011), 54.

[250] C. Leitherer, D. Schaerer, J.D. Goldader, R.M.G. Delgado, C. Robert, D.F. Kune, D.F. de Mello, D. Devost und T.M. Heckman, *ApJS* **123** (1999), 3.

[251] A. Lewis, A. Challinor und A. Lasenby, *ApJ* **538** (2000), 473.

[252] A. Lidz, C.A. Faucher-Giguère, A. Dall'Aglio, M. McQuinn, C. Fechner, M. Zaldarriaga, L. Hernquist und S. Dutta, *ApJ* **718** (2010), 199.

[253] A. Lidz und M. Malloy, *ApJ* **788** (2014), 175.

[254] A. Lidz, M. McQuinn, M. Zaldarriaga, L. Hernquist und S. Dutta, *ApJ* **670** (2007), 39.

[255] A. Loeb und S.R. Furlanetto: *The First Galaxies in the Universe*, 2013.

[256] M.S. Longair: *Theoretical Concepts in Physics*, Januar 2004.

[257] K. Lundmark, *MNRAS* **84** (1924), 747.

[258] C.P. Ma und E. Bertschinger, *ApJ* **455** (1995), 7.

[259] X. Ma, D. Kasen, P.F. Hopkins, C.A. Faucher-Giguère, E. Quataert, D. Kereš und N. Murray, *MNRAS* **453** (2015), 960.

[260] P. Madau und F. Haardt, *ArXiv e-prints* (2015).

[261] P. Madau, F. Haardt und M.J. Rees, *ApJ* **514** (1999), 648.

[262] P. Madau, A. Meiksin und M.J. Rees, *ApJ* **475** (1997), 429.

[263] P. Madau, L. Pozzetti und M. Dickinson, *ApJ* **498** (1998), 106.

[264] S. Majumdar, S. Bharadwaj, K.K. Datta und T.R. Choudhury, *MNRAS* **413** (2011), 1409.

[265] S. Malhotra und J.E. Rhoads, *ApJ* **617** (2004), L5.

[266] M. Malloy und A. Lidz, *ApJ* **767** (2013), 68.

[267] M. Malloy und A. Lidz, *ApJ* **799** (2015), 179.

[268] P. Martini, *Coevolution of Black Holes and Galaxies* (2004), 169.

[269] A. Maselli, B. Ciardi und A. Kanekar, *MNRAS* **393** (2009), 171.

[270] A. Maselli, A. Ferrara und B. Ciardi, *MNRAS* **345** (2003), 379.

[271] J.C. Mather, E.S. Cheng, R.E. Eplee, Jr. und et al., *ApJ* **354** (1990), L37.

[272] Y. Matsuoka, M. Onoue, N. Kashikawa und et al., *ArXiv e-prints* (2016).

[273] P. McDonald, J. Miralda-Escudé, M. Rauch, W.L.W. Sargent, T.A. Barlow und R. Cen, *ApJ* **562** (2001), 52.

[274] P. McDonald, J. Miralda-Escudé, M. Rauch, W.L.W. Sargent, T.A. Barlow, R. Cen und J.P. Ostriker, *ApJ* **543** (2000), 1.

[275] P. McDonald, U. Seljak, S. Burles und et al., *ApJS* **163** (2006), 80.

[276] P. McDonald, U. Seljak, R. Cen, P. Bode und J.P. Ostriker, *MNRAS* **360** (2005), 1471.

[277] I.D. McGreer, A. Mesinger und V. D'Odorico, *MNRAS* **447** (2015), 499.

[278] I.D. McGreer, A. Mesinger und X. Fan, *MNRAS* **415** (2011), 3237.

[279] C.F. McKee und E.C. Ostriker, *ARA&A* **45** (2007), 565.

[280] R.J. McLure, M. Cirasuolo, J.S. Dunlop, S. Foucaud und O. Almaini, *MNRAS* **395** (2009), 2196.

[281] R.J. McLure, J.S. Dunlop, L. de Ravel, M. Cirasuolo, R.S. Ellis, M. Schenker, B.E. Robertson, A.M. Koekemoer, D.P. Stark und R.A.A. Bowler, *MNRAS* **418** (2011), 2074.

[282] M. McQuinn, *ArXiv e-prints* (2015).

[283] M. McQuinn, L. Hernquist, A. Lidz und M. Zaldarriaga, *MNRAS* **415** (2011), 977.

[284] M. McQuinn, L. Hernquist, M. Zaldarriaga und S. Dutta, *MNRAS* **381** (2007), 75.

[285] M. McQuinn, A. Lidz, M. Zaldarriaga, L. Hernquist und S. Dutta, *MNRAS* **388** (2008), 1101.

[286] M. McQuinn, A. Lidz, M. Zaldarriaga, L. Hernquist, P.F. Hopkins, S. Dutta und C.A. Faucher-Giguère, *ApJ* **694** (2009), 842.

[287] M. McQuinn und E.R. Switzer, *Phys. Rev. D* **80** (2009), 063010.

[288] M. McQuinn und P.R. Upton Sanderbeck, *MNRAS* **456** (2016), 47.

[289] M. McQuinn und G. Worseck, *MNRAS* **440** (2014), 2406.

[290] A. Meiksin, *MNRAS* **356** (2005), 596.

[291] A. Meiksin und P. Madau, *ApJ* **412** (1993), 34.

[292] A. Meiksin, E.R. Tittley und C.K. Brown, *MNRAS* **401** (2010), 77.

[293] A. Meiksin und M. White, *MNRAS* **342** (2003), 1205.

[294] A. Meiksin und M. White, *MNRAS* **350** (2004), 1107.

[295] A.A. Meiksin, *Reviews of Modern Physics* **81** (2009), 1405.

[296] A. Melandri, M.G. Bernardini, P. D'Avanzo und et al., *A&Ap* **581** (2015), A86.

[297] G. Mellema, I.T. Iliev, M.A. Alvarez und P.R. Shapiro, *New Astro.* **11** (2006), 374.

[298] L.V.E. Mellema, G. Koopmans, F.A. Abdalla und et al., *Experimental Astronomy* **36** (2013), 235.

[299] A. Mesinger, *MNRAS* **407** (2010), 1328.

[300] A. Mesinger, A. Aykutalp, E. Vanzella, L. Pentericci, A. Ferrara und M. Dijkstra, *MNRAS* **446** (2015), 566.

[301] A. Mesinger und S. Furlanetto, *ApJ* **669** (2007), 663.

[302] A. Mesinger und S.R. Furlanetto, *MNRAS* **385** (2008), 1348.

[303] A. Mesinger und Z. Haiman, *ApJ* **611** (2004), L69.

[304] A. Mesinger und Z. Haiman, *ApJ* **660** (2007), 923.

[305] A. Mesinger, M. McQuinn und D.N. Spergel, *MNRAS* **422** (2012), 1403.

[306] D. Mihalas und B.W. Mihalas: *Foundations of radiation hydrodynamics*, 1984.

[307] J. Miralda-Escude, *ApJ* **501** (1998), 15.

[308] J. Miralda-Escudé, R. Cen, J.P. Ostriker und M. Rauch, *ApJ* **471** (1996), 582.

[309] J. Miralda-Escudé, M. Haehnelt und M.J. Rees, *ApJ* **530** (2000), 1.

[310] J. Miralda-Escudé und M.J. Rees, *MNRAS* **266** (1994), 343.

[311] H. Mo, F.C. van den Bosch und S. White: *Galaxy Formation and Evolution*, Mai 2010.

[312] M.F. Morales und J.S.B. Wyithe, *ARA&A* **48** (2010), 127.

[313] D.J. Mortlock, S.J. Warren, B.P. Venemans, M. Patel, P.C. Hewett, R.G. McMahon, C. Simpson, T. Theuns, E.A. Gonzáles-Solares, A. Adamson, S. Dye, N.C. Hambly, P. Hirst, M.J. Irwin, E. Kuiper, A. Lawrence und H.J.A. Röttgering, *Nature* **474** (2011), 616.

[314] R.E. Mostardi, A.E. Shapley, C.C. Steidel, R.F. Trainor, N.A. Reddy und B. Siana, *ApJ* **810** (2015), 107.

[315] A. Muzzin, D. Marchesini, M. Stefanon, M. Franx, H.J. McCracken, B. Milvang-Jensen, J.S. Dunlop, J.P.U. Fynbo, G. Brammer, I. Labbé und P.G. van Dokkum, *ApJ* **777** (2013), 18.

[316] K.K. Nilsson, O. Möller-Nilsson, P. Møller, J.P.U. Fynbo und A.E. Shapley, *MNRAS* **400** (2009), 232.

[317] M.L. Norman, D.R. Reynolds, G.C. So, R.P. Harkness und J.H. Wise, *ApJS* **216** (2015), 16.

[318] P. Noterdaeme, P. Petitjean, W.C. Carithers, I. Pâris, A. Font-Ribera, S. Bailey, E. Aubourg, D. Bizyaev, G. Ebelke, H. Finley, J. Ge, E. Malanushenko, V. Malanushenko, J. Miralda-Escudé, A.D. Myers, D. Oravetz, K. Pan, M.M. Pieri, N.P. Ross, D.P. Schneider, A. Simmons und D.G. York, *A&Ap* **547** (2012), L1.

[319] P.A. Oesch, R.J. Bouwens, G.D. Illingworth, I. Labbé, M. Franx, P.G. van Dokkum, M. Trenti, M. Stiavelli, V. Gonzalez und D. Magee, *ApJ* **773** (2013), 75.

[320] J.M. O'Meara, J.X. Prochaska, G. Worseck, H.W. Chen und P. Madau, *ApJ* **765** (2013), 137.

[321] Y. Ono, M. Ouchi, B. Mobasher und et al., *ApJ* **744** (2012), 83.

[322] D.E. Osterbrock und G.J. Ferland: *Astrophysics of gaseous nebulae and active galactic nuclei*, 2006.

[323] K. Ota, M. Iye, N. Kashikawa und et al., *ApJ* **677** (2008), 12.

[324] M. Ouchi, K. Shimasaku, M. Akiyama und et al., *ApJS* **176** (2008), 301.

[325] M. Ouchi, K. Shimasaku, H. Furusawa und et al., *ApJ* **723** (2010), 869.

[326] R.A. Overzier, R.J. Bouwens, G.D. Illingworth und M. Franx, *ApJ* **648** (2006), L5.

[327] J.P. Paardekooper, S. Khochfar und C. Dalla Vecchia, *MNRAS* **451** (2015), 2544.

[328] G. Paciga, J.G. Albert, K. Bandura, T.C. Chang, Y. Gupta, C. Hirata, J. Odegova, U.L. Pen, J.B. Peterson, J. Roy, J.R. Shaw, K. Sigurdson und T. Voytek, *MNRAS* **433** (2013), 639.

[329] G. Paciga, T.C. Chang, Y. Gupta, R. Nityanada, J. Odegova, U.L. Pen, J.B. Peterson, J. Roy und K. Sigurdson, *MNRAS* **413** (2011), 1174.

[330] N. Palanque-Delabrouille, C. Yèche, A. Borde und et al., *A&Ap* **559** (2013), A85.

[331] F. Paresce, C.F. McKee und S. Bowyer, *ApJ* **240** (1980), 387.

[332] A.R. Parsons, A. Liu, J.E. Aguirre, Z.S. Ali, R.F. Bradley, C.L. Carilli, D.R. DeBoer, M.R. Dexter, N.E. Gugliucci, D.C. Jacobs, P. Klima, D.H.E. MacMahon, J.R. Manley, D.F. Moore, J.C. Pober, I.I. Stefan und W.P. Walbrugh, *ArXiv e-prints* (2013).

[333] A.R. Parsons, A. Liu, J.E. Aguirre, Z.S. Ali, R.F. Bradley, C.L. Carilli, D.R. DeBoer, M.R. Dexter, N.E. Gugliucci, D.C. Jacobs, P. Klima, D.H.E. MacMahon, J.R. Manley, D.F. Moore, J.C. Pober, I.I. Stefan und W.P. Walbrugh, *ApJ* **788** (2014), 106.

[334] P. Paschos und M.L. Norman, *ApJ* **631** (2005), 59.

[335] M. Patel, S.J. Warren, D.J. Mortlock und J.P.U. Fynbo, *A&Ap* **512** (2010), L3.

[336] A.H. Pawlik und J. Schaye, *MNRAS* **389** (2008), 651.

[337] A.H. Pawlik, J. Schaye und C. Dalla Vecchia, *MNRAS* **451** (2015), 1586.

[338] A.H. Pawlik, J. Schaye und E. van Scherpenzeel, *MNRAS* **394** (2009), 1812.

[339] J.A. Peacock: *Cosmological Physics*, Januar 1999.

[340] J.A. Peacock: *Slipher, Galaxies, and Cosmological Velocity Fields. Slipher, Galaxies, and Cosmological Velocity Fields*, In *Origins of the Expanding Universe: 1912-1932*, herausgegeben von M. J. Way und D. Hunter, Band 471 von *Astronomical Society of the Pacific Conference Series*. (April 2013) Seite 3.

[341] J.A. Peacock, S. Cole, P. Norberg und et al., *Nature* **410** (2001), 169.

[342] J.A. Peacock und R.E. Smith, *MNRAS* **318** (2000), 1144.

[343] L. Pentericci, E. Vanzella, A. Fontana und et al., *ApJ* **793** (2014), 113.

[344] A.A. Penzias und R.W. Wilson, *ApJ* **142** (1965), 419.

[345] W.J. Percival, *ArXiv e-prints* (2013).

[346] W.J. Percival, C.M. Baugh, J. Bland-Hawthorn und et al., *MNRAS* **327** (2001), 1297.

[347] C. Péroux, R.G. McMahon, L.J. Storrie-Lombardi und M.J. Irwin, *MNRAS* **346** (2003), 1103.

[348] Planck Collaboration, R. Adam, P.A.R. Ade, N. Aghanim, Y. Akrami, M.I.R. Alves, M. Arnaud, F. Arroja, J. Aumont, C. Baccigalupi und et al., *ArXiv e-prints* (2015).

[349] Planck Collaboration, P.A.R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A.J. Banday und et al., *A&Ap* **571** (2014), A16.

[350] Planck Collaboration, P.A.R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A.J. Banday und et al., *A&Ap* **571** (2014), A17.

[351] Planck Collaboration, P.A.R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A.J. Banday, R.B. Barreiro, J.G. Bartlett und et al., *ArXiv e-prints* (2015).

[352] A. Pontzen, *Phys. Rev. D* **89** (2014), 083010.

[353] W.H. Press und P. Schechter, *ApJ* **187** (1974), 425.

[354] J.R. Pritchard und A. Loeb, *Reports on Progress in Physics* **75** (2012), 086901.

[355] J.X. Prochaska, S. Herbert-Fort und A.M. Wolfe, *ApJ* **635** (2005), 123.

[356] J.X. Prochaska und A.M. Wolfe, *ApJ* **696** (2009), 1543.

[357] E. Puchwein, J.S. Bolton, M.G. Haehnelt, P. Madau, G.D. Becker und F. Haardt, *MNRAS* **450** (2015), 4081.

[358] R. Quadri, P. van Dokkum, E. Gawiser, M. Franx, D. Marchesini, P. Lira, G. Rudnick, D. Herrera, J. Maza, M. Kriek, I. Labbé und H. Francke, *ApJ* **654** (2007), 138.

[359] A. Rahmati, A.H. Pawlik, M. Raicevic und J. Schaye, *MNRAS* **430** (2013), 2427.

[360] A. Rahmati und J. Schaye, *MNRAS* **438** (2014), 529.

[361] O. Rakic, J. Schaye, C.C. Steidel und G.C. Rudie, *MNRAS* **414** (2011), 3265.

[362] O. Rakic, J. Schaye, C.C. Steidel und G.C. Rudie, *ApJ* **751** (2012), 94.

[363] S. Raskutti, J.S. Bolton, J.S.B. Wyithe und G.D. Becker, *MNRAS* **421** (2012), 1969.

[364] M. Rauch, *ARA&A* **36** (1998), 267.

[365] M. Rauch, J. Miralda-Escudé, W.L.W. Sargent, T.A. Barlow, D.H. Weinberg, L. Hernquist, N. Katz, R. Cen und J.P. Ostriker, *ApJ* **489** (1997), 7.

[366] A.O. Razoumov, M.L. Norman, T. Abel und D. Scott, *ApJ* **572** (2002), 695.

[367] S.L. Reed, R.G. McMahon, M. Banerji, G.D. Becker und et al., *MNRAS* **454** (2015), 3952.

[368] M.J. Rees, *ARA&A* **22** (1984), 471.

[369] C.L. Reichardt: *Observing the Epoch of Reionization with the Cosmic Microwave Background. Observing the Epoch of Reionization with the Cosmic Microwave Background*, In *Astrophysics and Space Science Library*, herausgegeben von A. Mesinger, Band 423 von *Astrophysics and Space Science Library*. (2016) Seite 227.

[370] D.R. Reynolds, J.C. Hayes, P. Paschos und M.L. Norman, *Journal of Computational Physics* **228** (2009), 6833.

[371] M. Ricotti, N.Y. Gnedin und J.M. Shull, *ApJ* **534** (2000), 41.

[372] A.G. Riess, L. Macri, S. Casertano, H. Lampeitl, H.C. Ferguson, A.V. Filippenko, S.W. Jha, W. Li und R. Chornock, *ApJ* **730** (2011), 119.

[373] B.E. Robertson, R.S. Ellis, S.R. Furlanetto und J.S. Dunlop, *ApJ* **802** (2015), L19.

[374] B.E. Robertson, S.R. Furlanetto, E. Schneider, S. Charlot, R.S. Ellis, D.P. Stark, R.J. McLure, J.S. Dunlop, A. Koekemoer, M.A. Schenker, M. Ouchi, Y. Ono, E. Curtis-Lake, A.B. Rogers, R.A.A. Bowler und M. Cirasuolo, *ApJ* **768** (2013), 71.

[375] G. Rodighiero, E. Daddi, I. Baronchelli und et al., *ApJ* **739** (2011), L40.

[376] J. Rosdahl, J. Blaizot, D. Aubert, T. Stranex und R. Teyssier, *MNRAS* **436** (2013), 2188.

[377] G.C. Rudie, C.C. Steidel und M. Pettini, *ApJ* **757** (2012), L30.

[378] G.C. Rudie, C.C. Steidel, R.F. Trainor, O. Rakic, M. Bogosavljević, M. Pettini, N. Reddy, A.E. Shapley, D.K. Erb und D.R. Law, *ApJ* **750** (2012), 67.

[379] E.V. Ryan-Weber, *MNRAS* **367** (2006), 1251.

[380] G.B. Rybicki und I.P. dell'Antonio, *ApJ* **427** (1994), 603.

[381] R. Salvaterra, M. Della Valle, S. Campana und et al., *Nature* **461** (2009), 1258.

[382] R. Salvaterra, F. Haardt und A. Ferrara, *MNRAS* **362** (2005), L50.

[383] R. Salvaterra, F. Haardt und M. Volonteri, *MNRAS* **374** (2007), 761.

[384] M.R. Santos, *MNRAS* **349** (2004), 1137.

[385] D. Schaerer: *Evolutionary Synthesis Models as a Tool and Guide Towards the First Galaxies. Evolutionary Synthesis Models as a Tool and Guide Towards the First Galaxies*, In *The First Galaxies*, herausgegeben von T. Wiklind, B. Mobasher und V. Bromm, Band 396 von *Astrophysics and Space Science Library*. (2013) Seite 345.

[386] D. Schaerer, *ArXiv e-prints* (2014).

[387] D. Schaerer, S. de Barros und P. Sklias, *A&Ap* **549** (2013), A4.

[388] J. Schaye, *ApJ* **559** (2001), 507.

[389] J. Schaye, R.A. Crain, R.G. Bower, M. Furlong, M. Schaller, T. Theuns, C. Dalla Vecchia, C.S. Frenk, I.G. McCarthy, J.C. Helly, A. Jenkins, Y.M. Rosas-Guevara, S.D.M. White, M. Baes, C.M. Booth, P. Camps, J.F. Navarro, Y. Qu, A. Rahmati, T. Sawala, P.A. Thomas und J. Trayford, *MNRAS* **446** (2015), 521.

[390] J. Schaye, C. Dalla Vecchia, C.M. Booth, R.P.C. Wiersma, T. Theuns, M.R. Haas, S. Bertone, A.R. Duffy, I.G. McCarthy und F. van de Voort, *MNRAS* **402** (2010), 1536.

[391] J. Schaye, T. Theuns, M. Rauch, G. Efstathiou und W.L.W. Sargent, *MNRAS* **318** (2000), 817.

[392] M.A. Schenker, R.S. Ellis, N.P. Konidaris und D.P. Stark, *ApJ* **795** (2014), 20.

[393] J. Schroeder, A. Mesinger und Z. Haiman, *MNRAS* **428** (2013), 3058.

[394] J. Scott, J. Bechtold, A. Dobrzycki und V.P. Kulkarni, *ApJS* **130** (2000), 67.

[395] N.Z. Scoville. *Evolution of star formation and gas*, Oktober 2013. Seite 491.

[396] S. Seager, D.D. Sasselov und D. Scott, *ApJ* **523** (1999), L1.

[397] U. Seljak und M. Zaldarriaga, *ApJ* **469** (1996), 437.

[398] P.R. Shapiro und M.L. Giroux, *ApJ* **321** (1987), L107.

[399] P.R. Shapiro, I.T. Iliev, M.A. Alvarez und E. Scannapieco, *ApJ* **648** (2006), 922.

[400] A.E. Shapley, C.C. Steidel, M. Pettini und K.L. Adelberger, *ApJ* **588** (2003), 65.

[401] Y. Shen, M.A. Strauss, M. Oguri, J.F. Hennawi, X. Fan, G.T. Richards, P.B. Hall, J.E. Gunn, D.P. Schneider, A.S. Szalay, A.R. Thakar, D.E. Vanden Berk, S.F. Anderson, N.A. Bahcall, A.J. Connolly und G.R. Knapp, *AJ* **133** (2007), 2222.

[402] B.D. Sherwin, J. Dunkley, S. Das und et al., *Physical Review Letters* **107** (2011), 021302.

[403] M.S. Shin, H. Trac und R. Cen, *ApJ* **681** (2008), 756.

[404] A.M. Shone, S.L. Morris, N. Crighton und R.J. Wilman, *MNRAS* **402** (2010), 2520.

[405] J.M. Shull, B.D. Smith und C.W. Danforth, *ApJ* **759** (2012), 23.

[406] J.M. Shull und M.E. van Steenberg, *ApJ* **298** (1985), 268.

[407] B. Siana, H.I. Teplitz, J. Colbert, H.C. Ferguson, M. Dickinson, T.M. Brown, C.J. Conselice, D.F. de Mello, J.P. Gardner, M. Giavalisco und F. Menanteau, *ApJ* **668** (2007), 62.

[408] J. Silk, A. Di Cintio und I. Dvorkin, *ArXiv e-prints* (2013).

[409] R.A. Simcoe, P.W. Sullivan, K.L. Cooksey, M.M. Kao, M.S. Matejek und A.J. Burgasser, *Nature* **492** (2012), 79.

[410] V.M. Slipher, *Proceedings of the American Philosophical Society* **56** (1917), 403.

[411] A. Slosar, V. Iršič, D. Kirkby und et al., *JCAP* **4** (2013), 026.

[412] G.F. Smoot, C.L. Bennett, A. Kogut und et al., *ApJ* **396** (1992), L1.

[413] E. Sobacchi und A. Mesinger, *MNRAS* **453** (2015), 1843.

[414] R.S. Somerville und R. Davé, *ARA&A* **53** (2015), 51.

[415] A. Songaila und L.L. Cowie, *AJ* **123** (2002), 2183.

[416] A. Songaila und L.L. Cowie, *ApJ* **721** (2010), 1448.

[417] D.N. Spergel, R. Bean, O. Doré und et al., *ApJS* **170** (2007), 377.

[418] D.N. Spergel, L. Verde, H.V. Peiris und et al., *ApJS* **148** (2003), 175.

[419] L. Spitzer: *Physical Processes in the Interstellar Medium*, Mai 1998.

[420] V. Springel, *MNRAS* **364** (2005), 1105.

[421] V. Springel, *ARA&A* **48** (2010), 391.

[422] V. Springel, C.S. Frenk und S.D.M. White, *Nature* **440** (2006), 1137.

[423] V. Springel und L. Hernquist, *MNRAS* **339** (2003), 312.

[424] V. Springel, S.D.M. White, A. Jenkins und et al., *Nature* **435** (2005), 629.

[425] D.P. Stark, R.S. Ellis, K. Chiu, M. Ouchi und A. Bunker, *MNRAS* **408** (2010), 1628.

[426] D.P. Stark, M.A. Schenker, R. Ellis, B. Robertson, R. McLure und J. Dunlop, *ApJ* **763** (2013), 129.

[427] C.C. Steidel, K.L. Adelberger, A.E. Shapley, M. Pettini, M. Dickinson und M. Giavalisco, *ApJ* **592** (2003), 728.

[428] C.C. Steidel, D.K. Erb, A.E. Shapley, M. Pettini, N. Reddy, M. Bogosavljević, G.C. Rudie und O. Rakic, *ApJ* **717** (2010), 289.

[429] G. Steigman, *Annual Review of Nuclear and Particle Science* **57** (2007), 463.

[430] M. Stiavelli, S.G. Djorgovski, C. Pavlovsky, C. Scarlata, D. Stern, A. Mahabal, D. Thompson, M. Dickinson, N. Panagia und G. Meylan, *ApJ* **622** (2005), L1.

[431] C.M.S. Straatman, I. Labbé, L.R. Spitler und et al., *ApJ* **783** (2014), L14.

[432] R.S. Sutherland und M.A. Dopita, *ApJS* **88** (1993), 253.

[433] D. Syphers, S.F. Anderson, W. Zheng, A. Meiksin, D. Haggard, D.P. Schneider und D.G. York, *ApJ* **726** (2011), 111.

[434] Y. Takeuchi, S. Zaroubi und N. Sugiyama, *MNRAS* **444** (2014), 2236.

[435] N.R. Tanvir, D.B. Fox, A.J. Levan und et al., *Nature* **461** (2009), 1254.

[436] L.A.M. Tasca, O. Le Fèvre, N.P. Hathi und et al., *A&Ap* **581** (2015), A54.

[437] A. Tasitsiomi, *ApJ* **645** (2006), 792.

[438] N. Tejos, S.L. Morris, C.W. Finn, N.H.M. Crighton, J. Bechtold, B.T. Jannuzi, J. Schaye, T. Theuns, G. Altay, O. Le Fèvre, E. Ryan-Weber und R. Davé, *MNRAS* **437** (2014), 2017.

[439] R. Teyssier, *A&Ap* **385** (2002), 337.

[440] R. Teyssier, *ARA&A* **53** (2015), 325.

[441] T. Theuns, A. Leonard, G. Efstathiou, F.R. Pearce und P.A. Thomas, *MNRAS* **301** (1998), 478.

[442] T. Theuns, J. Schaye, S. Zaroubi, T.S. Kim, P. Tzanavaris und B. Carswell, *ApJ* **567** (2002), L103.

[443] T. Theuns, S. Zaroubi, T.S. Kim, P. Tzanavaris und R.F. Carswell, *MNRAS* **332** (2002), 367.

[444] R.M. Thomas und S. Zaroubi, *MNRAS* **384** (2008), 1080.

[445] V. Tilvi, C. Papovich, S.L. Finkelstein, J. Long, M. Song, M. Dickinson, H.C. Ferguson, A.M. Koekemoer, M. Giavalisco und B. Mobasher, *ApJ* **794** (2014), 5.

[446] B.M. Tinsley, *Fundamentals of Cosmic Physics* **5** (1980), 287.

[447] E.R. Tittley und A. Meiksin, *MNRAS* **380** (2007), 1369.

[448] A.R. Tomczak, R.F. Quadri, K.V.H. Tran und et al., *ApJ* **783** (2014), 85.

[449] T. Totani, K. Aoki, T. Hattori und et al., *PASJ* **66** (2014), 63.

[450] T. Totani, K. Aoki, T. Hattori und N. Kawai, *PASJ* **68** (2016), 15.

[451] T. Totani, N. Kawai, G. Kosugi, K. Aoki, T. Yamada, M. Iye, K. Ohta und T. Hattori, *PASJ* **58** (2006), 485.

[452] T. Treu, K.B. Schmidt, M. Trenti, L.D. Bradley und M. Stiavelli, *ApJ* **775** (2013), L29.

[453] T. Treu, M. Trenti, M. Stiavelli, M.W. Auger und L.D. Bradley, *ApJ* **747** (2012), 27.

[454] M. Trodden und S.M. Carroll, *ArXiv Astrophysics e-prints* (2004).

[455] P. Tummuangpak, R.M. Bielby, T. Shanks, T. Theuns, N.H.M. Crighton, H. Francke und L. Infante, *MNRAS* **442** (2014), 2094.

[456] M.L. Turner, J. Schaye, C.C. Steidel, G.C. Rudie und A.L. Strom, *MNRAS* **445** (2014), 794.

[457] M.L. Turner, J. Schaye, C.C. Steidel, G.C. Rudie und A.L. Strom, *MNRAS* **450** (2015), 2067.

[458] Y. Ueda, M. Akiyama, K. Ohta und T. Miyaji, *ApJ* **598** (2003), 886.

[459] Y. Utsumi, T. Goto, N. Kashikawa und et al., *ApJ* **721** (2010), 1680.

[460] M. Valdés und A. Ferrara, *MNRAS* **387** (2008), L8.

[461] F. van de Voort, J. Schaye, C.M. Booth, M.R. Haas und C. Dalla Vecchia, *MNRAS* **414** (2011), 2458.

[462] M.P. van Haarlem, M.W. Wise, A.W. Gunst und et al., *A&Ap* **556** (2013), A2.

[463] E. Vanzella, M. Giavalisco, M. Dickinson, S. Cristiani, M. Nonino, H. Kuntschner, P. Popesso, P. Rosati, A. Renzini, D. Stern, C. Cesarsky, H.C. Ferguson und R.A.E. Fosbury, *ApJ* **695** (2009), 1163.

[464] E. Vanzella, M. Giavalisco, A.K. Inoue, M. Nonino, F. Fontanot, S. Cristiani, A. Grazian, M. Dickinson, D. Stern, P. Tozzi, E. Giallongo, H. Ferguson, H. Spinrad, K. Boutsia, A. Fontana, P. Rosati und L. Pentericci, *ApJ* **725** (2010), 1011.

[465] A. Verhamme, D. Schaerer, H. Atek und C. Tapken, *A&Ap* **491** (2008), 89.

[466] A. Verhamme, D. Schaerer und A. Maselli, *A&Ap* **460** (2006), 397.

[467] M. Viel, J.S. Bolton und M.G. Haehnelt, *MNRAS* **399** (2009), L39.

[468] M. Viel, J. Schaye und C.M. Booth, *MNRAS* **429** (2013), 1734.

[469] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, S. Bird, D. Nelson und L. Hernquist, *Nature* **509** (2014), 177.

[470] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson und L. Hernquist, *MNRAS* **444** (2014), 1518.

[471] M. Volonteri, *The Astronomy and Astrophysics Review* **18** (2010), 279.

[472] J.W. Wadsley, J. Stadel und T. Quinn, *New Astro.* **9** (2004), 137.

[473] D.H. Weinberg, M.J. Mortonson, D.J. Eisenstein, C. Hirata, A.G. Riess und E. Rozo, *Phys. Rep.* **530** (2013), 87.

[474] R.L. White, R.H. Becker, X. Fan und M.A. Strauss, *AJ* **126** (2003), 1.

[475] S.D.M. White und C.S. Frenk, *ApJ* **379** (1991), 52.

[476] S.D.M. White und M.J. Rees, *MNRAS* **183** (1978), 341.

[477] S.D.M. White und V. Springel: *Where Are the First Stars Now? Where Are the First Stars Now?*, In *The First Stars*, herausgegeben von A. Weiss, T. G. Abel und V. Hill. (2000) Seite 327.

[478] J.A. Willick, *ArXiv Astrophysics e-prints* (1996).

[479] C.J. Willott, C.L. Carilli, J. Wagg und R. Wang, *ApJ* **807** (2015), 180.

[480] R.J. Wilman, S.L. Morris, B.T. Jannuzi, R. Davé und A.M. Shone, *MNRAS* **375** (2007), 735.

[481] J.H. Wise und T. Abel, *MNRAS* **414** (2011), 3458.

[482] J.H. Wise und R. Cen, *ApJ* **693** (2009), 984.

[483] A.M. Wolfe, E. Gawiser und J.X. Prochaska, *ARA&A* **43** (2005), 861.

[484] G. Worseck, J.X. Prochaska, M. McQuinn, A. Dall'Aglio, C. Fechner, J.F. Hennawi, D. Reimers, P. Richter und L. Wisotzki, *ApJ* **733** (2011), L24.

[485] S.A. Wouthuysen, *Physica* **18** (1952), 75.

[486] J.S.B. Wyithe und A. Loeb, *ApJ* **586** (2003), 693.

[487] J.S.B. Wyithe und A. Loeb, *ApJ* **610** (2004), 117.

[488] J.S.B. Wyithe, A. Loeb und C. Carilli, *ApJ* **628** (2005), 575.

[489] S. Wyithe und A. Loeb, *ArXiv Astrophysics e-prints* (2004).

[490] S. Yatawatta, A.G. de Bruyn, M.A. Brentjens und et al., *A&Ap* **550** (2013), A136.

[491] D.G. York, J. Adelman, J.E. Anderson, Jr. und et al., *AJ* **120** (2000), 1579.

[492] Q. Yu, *ApJ* **623** (2005), 683.

[493] Q. Yu und Y. Lu, *ApJ* **620** (2005), 31.

[494] O. Zahn, C.L. Reichardt, L. Shaw, A. Lidz und et al., *ApJ* **756** (2012), 65.

[495] M. Zaldarriaga, L. Hui und M. Tegmark, *ApJ* **557** (2001), 519.

[496] M. Zaldarriaga, D.N. Spergel und U. Seljak, *ApJ* **488** (1997), 1.

[497] S. Zaroubi, A.G. de Bruyn, G. Harker und R.M.e.a. Thomas, *MNRAS* **425** (2012), 2964.

[498] S. Zaroubi und J. Silk, *MNRAS* **360** (2005), L64.

[499] S. Zaroubi, R.M. Thomas, N. Sugiyama und J. Silk, *MNRAS* **375** (2007), 1269.

[500] Y. Zhang, A. Meiksin, P. Anninos und M.L. Norman, *ApJ* **495** (1998), 63.

[501] Z. Zheng, R. Cen, H. Trac und J. Miralda-Escudé, *ApJ* **716** (2010), 574.

[502] Z. Zheng, R. Cen, H. Trac und J. Miralda-Escudé, *ApJ* **726** (2011), 38.

Acknowledgments

I would like to acknowledge many people, without their help the thesis could not have been completed. Firstly, I would like to thank my advisor Benedetta Ciardi, who patiently allowed me to freely explore various aspects of reionization, Lyα radiative transfer, and 21cm cosmology. I would also like to thank Mark Dijkstra with many encouraging and insightful discussions. His deep insights on the nature of Lyα emitting galaxies and the physics of Lyα radiative transfer have deepened my understanding without doubt. Many thanks to Luca Graziani, who helped me how to run a radiative transfer code, CRASH, and the practical aspects of numerical simulations. Andy Chung and Michele Sasdelli helped me so many aspects on writing and coding issues, to name a few. They are the guys behind the scene, who helped me a lot throughout my PhD. My friends at MPA – Durand D'Souza, Haakon Andresen, Titouan Lazeyras, Fabian Schmidt, Inh Jee, Marius Eide – who made my PhD life pleasant and kept me going. Finally, I thanks to the two MPA directors, Simon White and Eiichiro Komatsu, allowing me to travel conferences and workshops. Last, but not the least, I thank to Kate O'Shea, Sonja Gruendl, and Gabi Kratschmann for their tremendous helps on housing issues and visa.

Finally, I cannot conclude my thesis without thanking my parents – Hideki and Atsuko Kakiichi – for allowing me this long journey from a little town in Japan, to Edinburgh, and to Munich for pursuing my dream to be an astrophysicist. Their long supports made all this possible.